Universidade Federal de Santa Catarina

BU

Forma de Jordan de uma Matriz e de um Operador

03738764

Graziela de Souza Sombrio

Março - 1999

TCC
UFSC
MTM
0091
Ex.1 BSCFM

Esta monografia foi julgada adequada como TRABALHO DE CONCLUSÃO DE CURSO no curso de Matemática - Habilitação Licenciatura, e aprovada em sua forma final pela banca de examinadores designados pela portaria $n^{\underline{o}}$ 01/SCG/99.

Prof. CARMEM SUZANE COMITRE GIMENEZ, Ms.

Professora de disciplina

Banca Examinadora:

Prof. Willian Glenn Whitley, Ph.D.

Orientador

Joana B.O. Quandt

Prof. Joana Benedita de Oliveira Quandt, Dra.

Prof. Licio Hernanes Bezerra, Dr.

Sumário

Introdução

Conceitos já estudados em Álgebra Linear foram estendidos para conjuntos lineares complexos. A partir disto, passamos a utilizar um outro tipo de matriz: as matrizes polinomiais. Depois de provarmos que tais matrizes podem ser transformadas em matrizes canônicas, passamos a estudar seus polinômios invariantes, bem como seus divisores elementares. Relações importantes foram realizadas envolvendo polinômios invariantes de matrizes polinomiais. Com isto conseguimos chegar à forma canônica de uma matriz. Não menos importante é a decomposição de espaços vetoriais em subespaços invariantes. Isto nos faz chegar ao estudo da forma canônica de uma matriz. Assim, conseguimos alcançar nosso maior objetivo: a Forma Canônica de Jordan.

Com isso, chegamos a uma relação de grande importância para o estudo de matrizes: *Toda matriz é semelhante a uma matriz de Jordan.*

Capítulo 1

Revisão

Neste capítulo iremos rever algumas definições, teoremas e propriedades importantes para o estudo de Álgebra Linear. Não faremos demonstrações, pois essas são vistas em disciplinas de Álgebra Linear e não temos como objetivo repetí-las neste trabalho.

Definição 1 *Dizemos que um conjunto $V \neq \emptyset$ é um espaço vetorial sobre $\mathbb{R}$ quando, e somente quando:*

1. *Existe uma adição $(u, v) \rightarrow u + v$ em V, com as seguintes propriedades:*

   (a) $u + v = v + u, \forall\ u, v \in V$;

   (b) $u + (v + w) = (u + v) + w, \forall\ u, v, w \in V$;

   (c) Existe em V um elemento neutro para essa adição o qual será simbolizado genericamente por $\mathbf{0}$. Ou seja:

   $$\exists\ \mathbf{0} \in V \mid u + \mathbf{0} = u, \forall\ u \in V;$$

   (d) Para todo elemento u de V existe o oposto, o qual indicaremos por $(-u)$. Assim:

   $$\forall\ u \in V, \exists\ (-u) \in V \mid u + (-u) = \mathbf{0}.$$

2. *Está definida uma multiplicação de $\mathbb{R} \times V$ em V, o que significa que a cada par (α, u) de $\mathbb{R} \times V$ está associado um único elemento de V que se indica por αu, e para essa multiplicação tem-se o seguinte:*

   (a) $\alpha(\beta u) = (\alpha\beta)u$;

   (b) $(\alpha + \beta)u = \alpha u + \beta u$;

(c) $\alpha(u+v) = \alpha u + \alpha v$;

(d) $1u = u$

para quaisquer u, v *de* V, α, β *de* $\mathbb{R}$.

Seja V um espaço vetorial sobre $\mathbb{R}$. São válidas as seguintes propriedades:

P 1 *Para todo* $\alpha \in \mathbb{R}, \alpha \mathbf{0} = \mathbf{0}$.

P 2 *Para todo* $u \in V, 0u = \mathbf{0}$.

P 3 *Uma igualdade* $\alpha u = \mathbf{0}$, *com* $\alpha \in \mathbb{R}$ *e* $u \in V$, *só é possível se* $\alpha = 0$ *ou* $u = \mathbf{0}$.

P 4 *Para todo* $\alpha \in \mathbb{R}$ *e todo* $u \in V$, $(-\alpha)u = \alpha(-u) = -\alpha u$.

P 5 *Quaisquer que sejam* $\alpha, \beta \in \mathbb{R}, u$ *e* $v \in V$, $(\alpha - \beta)u = \alpha u - \beta u$.

P 6 *Quaisquer que sejam* $\alpha \in \mathbb{R}$, u *e* $v \in V$, $\alpha(u - v) = \alpha u - \alpha v$.

P 7 *Dados* $\beta, \alpha_1, \ldots, \alpha_n$ *em* $\mathbb{R}$ *e* $u_1, \ldots, u_n$ *em* V, *então:*

$$\beta\left(\sum_{j=1}^{n} \alpha_j u_j\right) = \sum_{j=1}^{n} (\beta \alpha_j) u_j.$$

Definição 2 *Seja* V *um espaço vetorial sobre* $\mathbb{R}$. *Um subespaço vetorial de* V *é um subconjunto* $W \subset V$, *tal que:*

1. $\mathbf{0} \in W$;
2. $\forall u, v \in W, u + v \in W$;
3. $\forall \alpha \in \mathbb{R}$ *e* $\forall u \in W, \alpha u \in W$.

Definição 3 *Seja* V *um espaço vetorial sobre* $\mathbb{R}$. *Dizemos que* $L \subset V$ $(L \neq \emptyset)$ *é linearmente dependente (L.D.) se, e somente se, existem* $u_1, \ldots, u_n \in L$ e $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ *tal que*

$$\alpha_1 u_1 + \ldots + \alpha_n u_n = 0$$

sem que os escalares α_i *sejam todos iguais ao número zero.*

Se L *não é L.D., dizemos que* L *é linearmente independente (L.I.)*

Seja V um espaço vetorial sobre $\mathbb{R}$. Temos as seguintes propriedades:

P 8 *Se um conjunto finito $L \subset V$ contém o vetor nulo, então esse conjunto é L.D.*

P 9 *Se $S = \{u\} \subset V$ e $u \neq \mathbf{0}$, então S é L.I.*

P 10 *Se $S = \{u_1, \dots u_n\} \subset V$ é L.D., então um dos seus vetores é combinação linear dos outros.*

P 11 *Se S_1 e S_2 são subconjuntos finitos e não vazios de V, se $S_1 \subset S_2$ e S_1 é L.D., então S_2 também é L.D.*

P 12 *Se S_1 e S_2 são subconjuntos finitos e não vazios de V, com $S_1 \subset S_2$ e S_2 é L.I., então S_1 também é L.I.*

Definição 4 *Fixando vetores $u_1, \dots, u_n$ em V, o conjunto S de todos os vetores de V que são combinação linear destes, é um subespaço vetorial. S é chamado de subespaço gerado por $u_1, \dots, u_n$ e usamos a notação*

$$S = [u_1, \dots, u_n].$$

Dizemos que V é finitamente gerado se existe $S \subset V$, S finito, de maneira que $V = [S]$.

P 13 *Se $S = \{u_1, \dots, u_n\}$ é L.I., e para um certo $u \in V$ tivermos $S \cup \{u\} = \{u_1, \dots, u_n, u\}$ L.D., então o vetor u é combinação linear dos vetores $u_1, \dots u_n$, isto é, $u \in [S]$, onde $[S]$ é o subespaço de V gerado por S.*

P 14 *Se $S = \{u_1, \dots u_j, \dots, u_n\}$ e $u_j \in [S - u_j]$ (u_j é combinação linear dos demais vetores de S) então $[S] = [S - \{u_j\}]$.*

Definição 5 *Seja V um espaço vetorial finitamente gerado. Uma base de V é um subconjunto finito $B \subset V$ para o qual as seguintes condições se verificam:*

1. $[B] = V$.

2. B é linearmente independente.

Proposição 1 *Todo espaço vetorial finitamente gerado admite uma base.*

Daqui em diante, assumiremos que todos os espaços vetoriais usados são finitamente gerados. Como nosso objetivo é estudar os autovalores e autovetores de uma transformação através de matrizes finitas , todos os espaços usados serão finitamente gerados. Informamos que quase todas as afirmações a seguir são verdadeiras para espaços em geral, algumas com suaves alterações, mas freqüentemente possuem demonstrações diferentes das demonstrações para espaços finitamente gerados.

Proposição 2 *Seja* $B = \{u_1, u_2, \ldots, u_n\}$ *uma base de um espaço vetorial* V. *Se* $u \in V$ *e ainda se*

$$u = \alpha_1 u_1 + \ldots + \alpha_i u_i + \ldots + \alpha_n u_n$$

com $\alpha_i \neq 0$, *então o conjunto* $C = \{u_1, \ldots, u_{i-1}, u, u_{i+1}, \ldots, u_n\}$ *também é uma base de* V.

Proposição 3 *Suponhamos que exista uma base de* V *com* n *vetores. Então se* $B = \{u_1, \ldots, u_n\} \subset V$ *é L.I. e possui* n *vetores,* B *é também uma base de* V.

Proposição 4 *Seja* V *um espaço vetorial sobre* $\mathbb{R}$, *de dimensão* n. *Todo subconjunto de* V *que seja L.I. tem no máximo* n *vetores.*

Teorema 1 *(Teorema da invariância) Seja* V *um espaço vetorial finitamente gerado. Então duas bases quaisquer de* V *têm o mesmo número de vetores.*

Definição 6 *Seja* V *um espaço vetorial finitamente gerado. Denomina-se dimensão de* V *(notação: dim* V*) o número de vetores de qualquer uma de suas bases. Diz-se também, neste caso, que* V *é um espaço de dimensão finita.*

Teorema 2 *(Teorema do completamento) Seja* V *um espaço vetorial de dimensão* $n \geq 1$. *Se* $\{u_1, \ldots, u_r\} \subset V$ *é um subconjunto L.I. com* r *vetores e* $r < n$, *então existem* $n - r$ *vetores,* $u_{r+1}, \ldots u_n \in V$, *de maneira que* $B = \{u_1, \ldots u_r, u_{r+1}, \ldots u_n\}$ *é uma base de* V.

Proposição 5 *Todo subespaço vetorial de um espaço vetorial finitamente gerado é também finitamente gerado.*

Proposição 6 *Seja* W *um subespaço vetorial de* V *de dimensão finita. Se* $\dim W = \dim V$, *então* $W = V$.

Proposição 7 *Seja* W *um espaço vetorial sobre* $\mathbb{R}$ *de dimensão finita. Se* U *e* V *são subespaços de* W, *então:*

$$\dim (U \cap V) + \dim (U + V) = \dim U + \dim V.$$

Definição 7 *Seja* V *um espaço vetorial de dimensão* n *e consideremos duas bases de* V: $B = \{u_1, \ldots, u_n\}$ *e* $C = \{v_1, \ldots, v_n\}$. *Então existe uma única família de escalares* α_{ij} *de maneira que*

$$v_1 = \alpha_{11}u_1 + \ldots + \alpha_{n1}u_n$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$v_n = \alpha_{1n}u_1 + \ldots + \alpha_{nn}u_n$$

ou simplesmente

$$v_j = \sum_{i=1}^{n} \alpha_{ij}u_i \qquad (j = 1, 2, \ldots, n).$$

A matriz quadrada de ordem n

$$P = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

chama-se matriz de mudança de base C *para a base* B.

Definição 8 *Sejam* U *e* V *espaços vetoriais sobre* $\mathbb{R}$. *Uma aplicação* $F : U \to V$ *é chamada de transformação linear de* U *em* V *se, e somente se,*

1. $F(u_1 + u_2) = F(u_1) + F(u_2), \forall\ u_1, u_2 \in U$;
2. $F(\alpha u) = \alpha F(u), \forall\ \alpha \in \mathbb{R}$ *e* $\forall u \in U$.

No caso em que $U = V$, uma transformação linear $F : U \to U$ é chamada também de operador linear.

Sejam U e V espaços vetoriais sobre $\mathbb{R}$ e consideremos uma transformação linear $F : U \to V$. Valem as seguintes propriedades para F:

P 15 $F(0) = 0$;

P 16 $F(-u) = -F(u), \forall\ u \in U$;

P 17 $F(u_1 - u_2) = F(u_1) - F(u_2), \forall\ u_1, u_2 \in U$;

Definição 9 *A imagem de uma transformação linear* $F : U \to V$ *é dada por* $Im(F) = \{F(u) \mid u \in U\}$.

P 18 *Se* W *é um subespaço de* U, *então a imagem de* W *por* F *é um subespaço de* V.

P 19 *Sendo $F : U \to V$ linear então*

$$F(\sum_{i=1}^{n} a_i u_i) = \sum_{i=1}^{n} a_i F(u_i).$$

Definição 10 *Sejam U e V espaços vetoriais sobre $\mathbb{R}$ e $F : U \to V$ uma transformação linear. Indica-se por $Ker(F)$ e denomina-se núcleo de F o seguinte subconjunto de U:*

$$Ker(F) = \{u \in U \mid F(u) = \mathbf{0}\}.$$

Proposição 8 *Seja $F : U \to V$ uma transformação linear. Então:*

1. *$Ker(F)$ é um subespaço vetorial de U;*

2. *A transformação linear F é injetora se, e somente se, $Ker(F) = \{\mathbf{0}\}$.*

Teorema 3 *(Teorema do Núcleo e da Imagem) Sejam U e V espaços vetoriais de dimensão finita sobre $\mathbb{R}$. Dada uma transformação linear $F : U \to V$, então*

$$dim\ U = dim\ Ker(F) + dim\ Im(F).$$

Corolário 3.1 *Sejam U e V espaços vetoriais sobre $\mathbb{R}$ com a mesma dimensão finita n e suponhamos que $F : U \to V$ seja uma transformação linear. Então são equivalentes as seguintes afirmações:*

1. *F é sobrejetora;*

2. *F é bijetora;*

3. *F é injetora;*

4. *F transforma uma base de U em uma base de V (ou seja, se B é uma base de U, então $F(B)$ é base de V).*

Definição 11 *Entende-se por isomorfismo do espaço vetorial U no espaço vetorial V uma transformação linear $F : U \to V$ que seja bijetora. Um isomorfismo $F : U \to U$ é um automorfismo de U.*

Proposição 9 *Se F é um isomorfismo de U em V, então $F^{-1} : V \to U$ também é um isomorfismo (de V em U).*

Lema 1 *Sejam U e V espaços vetoriais sobre $\mathbb{R}$. Se dim $U = n$ e $B = \{u_1, u_2, \ldots, u_n\}$ é uma base de U, então para toda seqüência $v_1, \ldots, v_n$ de vetores de V, a aplicação $F : U \to V$, definida por*

$$F(\sum_{i=1}^{n} \alpha_i u_i) = \sum_{i=1}^{n} \alpha_i v_i$$

é linear e $F(u_i) = v_i (i = 1, 2, \ldots, n)$. Ademais, se $G : U \to V$ é linear e $G(u_i) = v_i (i = 1, 2, \ldots, n)$, então $G = F$.

Teorema 4 *Dois espaços U e V de dimensão finita são isomorfos se, e somente se, $\dim U = \dim V$.*

Sejam U e V espaços vetoriais sobre $\mathbb{R}$. Indicaremos por $L(U, V)$ o conjunto das transformações lineares de U em V.

Definição 12 *Dados $F, G \in L(U, V)$, definimos a soma $F + G$ de F com G da seguinte maneira:*

$$F + G : U \to V \text{ e } (F + G)(u) = F(u) + G(u), \forall u \in U.$$

Sejam $(F, G) \to F + G$ em $L(U, V)$. Valem as seguintes propriedades:

1. Associativa: $F + (G + H) = (F + G) + H, \forall\ F, G, H \in L(U, V)$;

2. Comutativa: $F + G = G + F, \forall\ F, G \in L(U, V)$;

3. Existe elemento neutro: a transformação linear nula $0 : U \to V$ é tal que $F + 0 = F, \forall\ F \in L(U, V)$;

4. Para toda transformação $F \in L(U, V)$ existe neste conjunto a transformação oposta:
$$\exists\ (-F) \in L(U, V) \mid F + (-F) = \mathbf{0}.$$

Definição 13 *Dados $F \in L(U, V)$ e $\alpha \in \mathbb{R}$, definimos o produto αF de F por α da seguinte forma:*

$$\alpha F : U \to V \text{ e } (\alpha F)(u) = \alpha F(u), \forall\ u \in U.$$

Aqui valem as seguintes propriedades:

1. $(\alpha\beta)F = \alpha(\beta F)$;

2. $(\alpha + \beta)F = \alpha F + \beta F$;

3. $\alpha(F + G) = \alpha F + \alpha G$;

4. $1F = F$;

quaisquer que sejam α e β em $\mathbb{R}$ e F e G em $L(U, V)$.

Definição 14 *Sejam U, V e W espaços vetoriais sobre $\mathbb{R}$. Se $F : U \to V$ e $G : V \to W$ são transformações lineares, define-se a aplicação composta de F e G (notação: $G \circ F$) da seguinte maneira:*

$$G \circ F : U \to W \text{ e } (G \circ F)(u) = G(F(u)), \forall\ u \in U.$$

Consideremos o caso $U = V = W$. Quando isto acontece $(G, F) \to G \circ F$ passa a ser uma operação em L(U) que apresenta as seguintes propriedades:

1. $(H \circ G) \circ F = H \circ (G \circ F), \forall\ H, G, F \in L(U)$;
2. $I \circ F = F \circ I = F, \forall\ F \in L(U)$;
3. $H \circ (F + G) = H \circ F + H \circ G$ e $(F + G) \circ H = (F \circ H) + (G \circ H)$; $\forall\ F, G, H \in L(U)$.

Nota: No conjunto L(U) define-se potenciação para expoentes naturais assim: $F^0 = I$ (operador idêntico); $F^1 = F$; $F^2 = F \circ F$; $F^3 = F \circ F \circ F$;... Contudo é bom observar que para essa potenciação podemos ter resultados em princípio curiosos como $F^2 = I$, com $F \neq I$ e $F \neq -I$, $F^n = 0$ (operador nulo) com $F \neq 0$. Um operador $F \in L(U)$ tal que $F^2 = F$ chama-se idempotente (ou projeção); se $F^n = 0$, para um certo número natural n, então F se diz nilpotente.

Sejam U e V espaços vetoriais de dimensão n e m, respectivamente, sobre $\mathbb{R}$. Consideremos uma transformação linear $F : U \to V$. Dadas as bases $B = \{u_1, \ldots, u_n\}$ de U e $C = \{v_1, \ldots, v_m\}$ de V, então cada um dos vetores $F(u_1), \ldots, F(u_n)$ está em V e conseqüentemente é combinação linear da base C:

$$F(u_1) = \alpha_{11}v_1 + \alpha_{21}v_2 + \ldots + \alpha_{m1}v_m$$

$$F(u_2) = \alpha_{12}v_1 + \alpha_{22}v_2 + \ldots + \alpha_{m2}v_m$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$F(u_n) = \alpha_{1n}v_1 + \alpha_{2n}v_2 + \ldots + \alpha_{mn}v_m$$

Definição 15 *A matriz $m \times n$ sobre $\mathbb{R}$*

$$(\alpha_{ij}) = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}$$

que se obtém da consideração anterior é chamada matriz de F em relação às bases B e C. Usaremos, para indicar essa matriz, a notação

$$[F]_C^B.$$

Toda matriz $m \times n$ está associada a uma transformação linear $T : \mathbb{R}^n \to \mathbb{R}^m$.

Sejam $\beta = \{v_1, \ldots, v_n\}$ uma base de $\mathbb{R}^n$, $\beta' = \{w_1, \ldots, w_m\}$ uma base de $\mathbb{R}^m$ e

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

Podemos associar

$$T_A : \mathbb{R}^n \to \mathbb{R}^m$$

$$v \to T_A(v)$$

Como segue:

Seja $X = [v]_\beta = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

$$AX = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

Então, $T_A(v) = y_1 w_1 + \ldots + y_m w_m$ onde $y_i = A_i X$ e A_i é a i-ésima linha de A.

Observe que T passa a ser a aplicação linear associada à matriz A e bases β e β', isto é $T = T_A$.

Proposição 10 *Sejam U e V espaços vetoriais sobre $\mathbb{R}$, de dimensões n e m, respectivamente. Então, fixadas as bases $B = \{u_1, \ldots, u_n\}$ e $C = \{v_1, \ldots, v_m\}$ de U e V, respectivamente, a aplicação $F \to (F)$ que a cada $F \in L(U, V)$ associa a matriz de F em relação às bases B e C é bijetora.*

Teorema 5 *Sejam V e W espaços vetoriais, α base de V, β base de W e $T : V \to W$ uma aplicação linear. Então para todo $v \in V$ vale:*

$$[T(v)]_\beta = [T]^\alpha_\beta . [v]_\alpha .$$

Teorema 6 *Seja $T : V \to W$ uma aplicação linear e α e β bases de V e W respectivamente. Então: dim $Im(T)$ = posto de $[T]^\alpha_\beta$; dim $Ker(T)$ = nulidade* [1] *de $[T]^\alpha_\beta$ = número de colunas $[T]^\alpha_\beta$− posto de $[T]^\alpha_\beta$.*

Teorema 7 *Sejam $T_1 : V \to W$ e $T_2 : W \to U$ transformações lineares e α, β, γ bases de V, W e U respectivamente. Então a composta de T_1 com T_2, $T_2 \circ T_1 : V \to U$, é linear e*

$$[T_2 \circ T_1]^\alpha_\gamma = [T_2]^\beta_\gamma . [T_1]^\alpha_\beta .$$

Corolário 7.1 *Se $T : V \to W$ é uma transformação linear inversível (T é um isomorfismo) e α e β são as bases de V e W, então $T^{-1} : W \to V$ é um operador linear e*

$$[T^{-1}]^\beta_\alpha = ([T]^\alpha_\beta)^{-1}.$$

Corolário 7.2 *Seja $T : V \to W$ uma transformação linear e α e β bases de V e W. Então T é inversível se e somente se* $\det[T]^\alpha_\beta \neq 0$.

Corolário 7.3

$$[T]^{\alpha'}_{\beta'} = [I \circ T \circ I]^{\alpha'}_{\beta'} = [I]^\beta_{\beta'} [T]^\alpha_\beta [I]^{\alpha'}_\alpha$$

$$\begin{tikzcd}
\alpha & & \beta \\
V \arrow[rr] & & W \arrow[d] \\
V \arrow[u] \arrow[rr] & & W \\
\alpha' & & \beta'
\end{tikzcd}$$

Definição 16 *Sejam P e Q duas matrizes quadradas e de ordem n. Dizemos que P é semelhante a Q se, e somente se, existe uma matriz inversível M, também de ordem n, de modo que:*

$$P = M^{-1} Q M.$$

[1] nulidade de uma matriz é o número de variáveis independentes, que podem assumir valores arbitrários

Duas matrizes correspondentes a um mesmo operador linear em $\mathbb{R}$ para bases distintas são semelhantes, ou seja, para um operador linear em $\mathbb{R}$ há classes de matrizes semelhantes correspondentes; elas representam o operador dado em várias bases.

Capítulo 2

Autovalores, autovetores e polinômio característico de uma matriz

2.1 Fórmula de Binet-Cauchy

Seja $C = (c_{ij})_1^m$ uma matriz quadrada tal que $C = AB$, onde $A = (a_{ik})$ possui dimensão $m \times n$ e $B = (b_{kj})$ possui dimensão $n \times m$. Então

$$\begin{pmatrix} c_{11} & \dots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mm} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nm} \end{pmatrix},$$

ou seja,

$$C_{ij} = \sum_{\alpha=1}^{n} a_{i\alpha} b_{\alpha j} \quad (i, j = 1, 2, \dots, m).$$

A fórmula de Binet-Cauchy expressa o determinante de C em termos dos menores de A e B:

$$\begin{vmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mm} \end{vmatrix} = \sum_{1 \le k_1 < k_2 < \cdots < k_m \le n} \begin{vmatrix} a_{1k_1} & \cdots & a_{1k_m} \\ \vdots & \ddots & \vdots \\ a_{mk_1} & \cdots & a_{mk_m} \end{vmatrix} \begin{vmatrix} b_{k_1 1} & \cdots & b_{k_1 m} \\ \vdots & \ddots & \vdots \\ b_{k_m 1} & \cdots & b_{k_m m} \end{vmatrix},$$

ou ainda,

$$C\begin{pmatrix} 1 & 2 & \cdots & m \\ 1 & 2 & \cdots & m \end{pmatrix} = \sum_{1 \le k_1 < k_2 < \cdots < k_m \le n} A\begin{pmatrix} 1 & 2 & \cdots & m \\ k_1 & k_2 & \cdots & k_m \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_m \\ 1 & 2 & \cdots & m \end{pmatrix}.$$

Como conseqüência da fórmula de Binet-Cauchy, podemos expressar os menores do produto de duas matrizes retangulares em termos dos menores dos fatores.

Sejam $A = (a_{ik})$, $B = (b_{ik})$ e $C = (c_{ik})$, onde $i = 1, 2, \ldots, m$, $k = 1, 2, \ldots, n$, $j = 1, 2, \ldots, q$ e $C = AB$.

Consideramos um menor arbitrário de C:
$C\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}$ tal que $\begin{pmatrix} 1 \leq i_1 \leq i_2 \leq \cdots \leq i_p \leq m \\ 1 \leq j_1 \leq j_2 \leq \cdots \leq j_p \leq q \end{pmatrix}$ onde $p \leq m$ e $p \leq q$.

Aplicando a fórmula de Binet-Cauchy temos:

$$C\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix} \sum_{1 \leq k_1 \leq k_2 \leq \cdots \leq k_p \leq m} A\begin{pmatrix} i_1 & i_2 & \cdots & i_p \\ k_1 & k_2 & \cdots & k_p \end{pmatrix} B\begin{pmatrix} k_1 & k_2 & \cdots & k_p \\ j_1 & j_2 & \cdots & j_p \end{pmatrix}.$$

2.2 Autovalores e autovetores

Definimos no capítulo anterior *Espaço vetorial sobre* $\mathbb{R}$. A partir deste capítulo iremos trabalhar com *Espaços vetoriais sobre* $\mathbb{C}$.

Definição 17 *Dizemos que um conjunto $V \neq \emptyset$ é um espaço vetorial sobre $\mathbb{C}$ quando e somente quando:*

1. *Existe uma adição $(u, v) \to u + v$ em V, com as seguintes propriedades:*

   (a) $u + v = v + u, \forall\ u, v \in V$;

   (b) $u + (v + w) = (u + v) + w, \forall\ u, v, w \in V$;

   (c) *Existe em V um elemento neutro para essa adição o qual será simbolizado genericamente por $\mathbf{0}$. Ou seja:*

   $$\exists\ \mathbf{0} \in V,\ |\ u + \mathbf{0} = u, \forall\ u \in V;$$

   (d) *Para todo elemento u de V existe o oposto; indicaremos por $(-u)$ esse oposto. Assim:*

   $$\forall\ u \in V, \exists\ (-u) \in V\ |\ u + (-u) = \mathbf{0}.$$

2. *Está definida uma multiplicação de $\mathbb{C} \times V$ em V, o que significa que a cada par (α, u) de $\mathbb{C} \times V$ está associado um único elemento de V que se indica por αu, e para essa multiplicação tem-se o seguinte:*

   (a) $\alpha(\beta u) = (\alpha\beta)u$;

(b) $(\alpha + \beta)u = \alpha u + \beta u$;

(c) $\alpha(u + v) = \alpha u + \alpha v$;

(d) $1u = u$

para quaisquer u, v de V, α, β de $\mathbb{C}$.

Podemos assumir como verdadeiras todas as afirmações citadas no capítulo anterior. Isto se deve ao fato de que as demonstrações utilizam as regras de aritmética dos escalares, as quais permanecem inalteradas quando passamos a trabalhar com números complexos. Para exemplificar este fato, iremos mostrar como foi feita a demonstração do Corolário 3.1 do Teorema 3 (Teorema do Núcleo e da Imagem) , que se encontra no Capítulo 1.

Corolário 3.1 *Sejam U e V espaços vetoriais sobre $\mathbb{R}$ com a mesma dimensão finita n e suponhamos que $F : U \to V$ é uma transformação linear. Então são equivalentes as seguintes afirmações:*

1. *F é sobrejetora;*

2. *F é bijetora;*

3. *F é injetora;*

4. *F transforma uma base de U em uma base de V (ou seja, se B é uma base de U, então $F(B)$ é base de V).*

Demonstração:

$(I) \Rightarrow (II)$ *Se F é sobrejetora então F é injetora.*

Por hipótese, $Im(F) = V$. Levando em conta que $dim\ U = dim\ V$, a fórmula $dim\ U = dim\ Ker(F) + dim\ Im(F)$ equivale então a $dim\ Ker(F) = 0$. Logo $Ker(F) = \{0\}$ e F é injetora. Então F é bijetora.

$(II) \Rightarrow (III)$ *Se F é bijetora então F é injetora.*

Essa demonstração é imediata pois, dada uma função F, dizemos que F é bijetora se e somente se, F é injetora e sobrejetora.

$(III) \Rightarrow (IV)$ *Se F é injetora então F transforma uma base de U em uma base de V.*

Sendo $B = \{u_1, \ldots, u_n\}$ uma base de U mostremos que $F(B) = \{f(u_1), \ldots, f(u_n)\}$ é uma base de V. Observamos de início que $F(B)$ tem a mesma quantidade de vetores que B, pois F é injetora. Então basta mostrar que $F(B)$ é L.I. . Suponhamos

$\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ (neste caso $\mathbb{C}$) e $\alpha_1 F(u_1) + \ldots + \alpha_n F(u_n) = \mathbf{0}$. Disto resulta, pela linearidade de F, que

$$F(\alpha_1 u_1 + \ldots + \alpha_n u_n) = \mathbf{0}.$$

Sendo F injetora segue que

$$\alpha_1 u_1 + \ldots + \alpha_n u_n = \mathbf{0}.$$

Como B é L.I., conclui-se que $\alpha_1 = \alpha_2 = \ldots = \alpha_n = \mathbf{0}$

$(IV) \Rightarrow (I)$ *Se F transforma uma base de U em uma base de V então F é sobrejetora.*

Seja $v \in V$. Tomando uma base $B = \{u_1, \ldots, u_n\}$ de U, então nossa hipótese garante que $F(B) = \{F(u_1), \ldots, F(u_n)\}$ é uma base de V. Logo v é combinação linear de $F(B)$:

$$v = \alpha_1 F(u_1) + \ldots + \alpha_n F(u_n), \quad \textit{com } \alpha_1, \ldots, \alpha_n \in \mathbb{R} \ (\textit{neste caso } \mathbb{C}).$$

Como F é linear podemos afirmar que

$$v = F(\alpha_1 u_1 + \ldots + \alpha_n u_n).$$

Estando em U a combinação linear $\alpha_1 u_1 + \ldots + \alpha_n u_n$ ficou provado que todo elemento de V é imagem, por F, de um elemento de U. Ou seja, F é sobrejetora. Isso completa a demonstração.

Como pode ser visto nesta demonstração retirada de [2] pg 115, as afirmações para espaços vetoriais complexos são quase idênticas às afirmações para espaços vetoriais reais, com algumas alterações triviais.

Nas discussões seguintes, denotaremos por $\mathbb{F}$ quaisquer um dos corpos de escalares $\mathbb{R}$ ou $\mathbb{C}$.

Dada uma transformação linear de um espaço vetorial $T : V \to V$, procuramos vetores que são levados em múltiplos de si mesmo, ou seja, queremos um vetor $v \in V$ e um escalar $\lambda \in \mathbb{F}$ tais que

$$T(v) = \lambda v.$$

Definição 18 *Seja $T : V \to V$ um operador linear. Se existirem $v \in V$, $v \neq 0$, e $\lambda \in \mathbb{F}$ tais que $T(v) = \lambda v$, λ é um autovalor de T e v um autovetor de T associado a λ.*

Seja A uma matriz quadrada de ordem n. Um autovalor de A é um autovetor da transformação linear $T_A : \mathbb{F}^n \to \mathbb{F}^n$ associada à matriz A em relação à base canônica, isto é, $T_A(v) = Av$. Assim, $\lambda \in \mathbb{F}$ é um autovalor de A se $\exists\, v(\neq 0) \in \mathbb{F}^n$ tal que $Av = \lambda v$; $v \in \mathbb{F}^n$ é um autovetor subordinado ao autovalor λ se $Av = \lambda v$

Vamos agora estudar um método prático para encontrarmos os autovalores e os autovetores da matriz A. Sabemos que tais autovalores e autovetores satisfazem a equação $Av = \lambda v$ ou $Av = (\lambda I)v$ ou ainda $(A - \lambda I)v = 0$. Escrevendo explicitamente esta equação temos:

$$\begin{pmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Se $\det(A - \lambda I) \neq 0$, temos que o rank desta matriz é n e portanto o sistema de equações lineares homogêneo tem uma única solução. Mas $x_1 = x_2 = \ldots = x_n = 0$ sempre é solução de um sistema homogêneo. Então a única solução possível seria a nula. Então para encontrarmos os autovalores v precisamos ter $\det(A - \lambda I) = 0$.

2.3 Polinômio característico

Definição 19 *Chamamos de matriz característica de A, a matriz $\lambda I - A$, onde $A = (a_{ik})_1^n$. O determinante da matriz característica*

$$\Delta(\lambda) = \det(\lambda I - A)$$

é um polinômio escalar em λ e é chamado de polinômio característico de A.

Definição 20 *Seja V um espaço vetorial de dimensão n e $T : V \to V$ um operador linear. Chama-se polinômio característico de T o polinômio característico da matriz de T em relação a qualquer base de V.*

Definição 21 *O subespaço $V_\lambda = \{v \in V : T(v) = \lambda v\}$ é chamado de subespaço associado ao autovalor λ.*

Quando $\mathbb{F} = \mathbb{R}$, este polinômio terá coeficientes reais mas, em alguns casos, poderá não ter raízes reais. Este fato inibe a relação de autovalores com as raízes da equação característica. Entretanto, se $\mathbb{F} = \mathbb{C}$, o fato deixa de existir.

Para o caso de um espaço vetorial complexo, todas as raízes do polinômio característico serão complexas.

Daqui em diante todos os espaços serão considerados complexos, salvo quando dito o contrário.

Definição 22 *Chamamos de multiplicidade algébrica de um autovalor a quantidade de vezes que ele aparece como raiz do polinômio característico e de multiplicidade geométrica de um autovalor a dimensão do subespaço V_λ de autovetores associados a λ.*

Seja $F(\lambda)$ um polinômio matricial de ordem n

$$F(\lambda) = F_0\lambda^m + F_1\lambda^{m-1} + \ldots + F_m \quad (F_0 \neq 0),$$

onde $F_0, \cdots, F_m$ são matrizes quadradas e de mesma ordem.

Podemos escrever este polinômio da seguinte maneira:

$$F(\lambda) = \lambda^m F_0 + \lambda^{m-1}F_1 + \ldots + F_m \quad (F_0 \neq 0).$$

Para qualquer escalar λ, os resultados das duas equações acima são iguais. Entretanto, se substituirmos λ por uma matriz quadrada A de ordem n, não podemos garantir que os resultados das duas equações permanecem iguais, pois as potências de A não são, necessariamente, permutáveis com os coeficientes matriciais F_0, F_1, ..., F_m.

Temos

$$F(A) = F_0A^m + F_1A^{m-1} + \ldots + F_m \quad (F_0 \neq 0)$$

e

$$\hat{F}(A) = A^m F_0 + A^{m-1}F_1 + \ldots + F_m \quad (F_0 \neq 0).$$

Chamamos $F(A)$ de valor direito e $\hat{F}(A)$ de valor esquerdo na substituição de A por λ. Valor direito significa que as potências de A estão à direita dos coeficientes e valor esquerdo significa que as potências de A estão à esquerda dos mesmos.

Teorema 8 *(Teorema de Bézout generalizado) Quando a matriz polinomial $F(\lambda)$ é dividida pela direita pelo binômio $\lambda I - A$, o resto é $F(A)$, quando é dividida pela esquerda, o resto é $\hat{F}(A)$.*

Demonstração:

Seja $F(\lambda)$ um polinômio matricial de ordem n.

$$F(\lambda) = F_0\lambda^m + F_1\lambda^{m-1} + \ldots + F_m \quad (F_0 \neq 0).$$

Dividimos $F(\lambda)$ por $\lambda I - A$. Neste caso o resto direito, $R(\lambda)$, e o resto esquerdo, $\hat{R}(\lambda)$, não dependem de λ. Para determinar o resto direito usamos a divisão usual:

$$F(\lambda) = F_0\lambda^m + F_1\lambda^{m-1} + \ldots + F_m \quad (F_0 \neq 0) =$$

$$= F_0\lambda^{m-1}(\lambda I - A) + (F_0A + F_1)\lambda^{m-1} + F_2\lambda^{m-2} + \ldots$$

$$= [F_0\lambda^{m-1} + (F_0A + F_1)\lambda^{m-2}](\lambda I - A) + (F_0A^2 + F_1A + F_2)\lambda^{m-2} + F_3\lambda^{m-3} + \ldots$$

$$= [F_0\lambda^{m-1} + (F_0A + F_1)\lambda^{m-2} + \ldots + (F_0A^{m-1} + F_1A^{m-2} + \ldots + F_{m-1})](\lambda I - A) +$$

$$+ F_0A^m + F_1A^{m-1} + \ldots + F_m$$

Temos então que

$$R = F_0A^m + F_1A^{m-1} + \ldots + F_m = F(A).$$

Da mesma forma encontramos o resto esquerdo da divisão de $F(\lambda)$ por $\lambda I - A$:

$$\hat{R} = A^mF_0 + A^{m-1}F_1 + \ldots + F_m = \hat{F}(A).$$

Isso completa a demonstração do teorema.

Definição 23 *Uma matriz polinomial, ou λ-matriz, é uma matriz retangular $A(\lambda)$, na qual seus elementos são polinômios em λ:*

$$A(\lambda) = (a_{ik}(\lambda)) = (a_{ik}^0\lambda^l + a_{ik}^1\lambda^{l-1} + \ldots + a_{ik}^l)(i = 1, 2, \ldots, m; k = 1, 2, \ldots, n);$$

onde l é o maior grau do polinômio $a_{ik}(\lambda)$.

Dada uma matriz A, dizemos que o cofator Δ_{ij} do elemento a_{ij} da matriz é $(-1)^{i+j}\det A_{ij}$, onde A_{ij} é a submatriz de A, obtida extraindo-se a i-ésima linha e a j-ésima coluna. Com estes cofatores podemos formar uma nova matriz A, denominada matriz dos cofatores de A.

Definição 24 *Dada uma matriz quadrada A, chamamos de matriz adjunta clássica da matriz $(\lambda I - A)$, a matriz transposta $B(\lambda)$ da matriz dos cofatores de $(\lambda I - A)$.*

Segue desta definição que:

$$(\lambda I - A)B(\lambda) = \Delta(\lambda)I$$

$$B(\lambda)(\lambda I - A) = \Delta(\lambda)I$$

Teorema 9 *(Teorema de Hamilton-Cayley) Toda matriz quadrada A satisfaz sua equação característica, isto é,*

$$P(A) = 0$$

Demonstração:

Seja A uma matriz quadrada de ordem n e $\Delta(\lambda)$ seu polinômio característico:

$$\Delta(\lambda) = \det(\lambda I - A) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$$

Chamamos de $B(\lambda)$ a matriz adjunta de $(\lambda I - A)$. Como os elementos de $B(\lambda)$ são cofatores de $(\lambda I - A)$, eles são polinômios de grau menor ou igual a $n-1$ em λ. Assim podemos escrever $B(\lambda)$ na forma:

$$B(\lambda) = B_{n-1}\lambda^{n-1} + \ldots + B_1\lambda + B_0.$$

Podemos aplicar a seguinte identidade:

$$(\lambda I - A)B(\lambda) = \Delta(\lambda)I,$$

ou ainda,

$$(\lambda I - A)(B_{n-1}\lambda^{n-1} + \ldots + B_1\lambda + B_0) = (\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0)I$$

$$\Rightarrow B_{n-1}\lambda^n + \ldots + B_1\lambda^2 + B_0\lambda - AB_{n-1}\lambda^{n-1} - \ldots - B_1A\lambda - B_0A =$$
$$= \lambda^n I + a_{n-1}\lambda^{n-1}I + \ldots + a_1\lambda I + a_0 I$$

Igualando os coeficientes das potências correspondentes de λ temos:

$$B_{n-1}\lambda^n = \lambda^n I \Rightarrow B_{n-1} = I$$

$$B_{n-2}\lambda^{n-1} - AB_{n-1}\lambda^{n-1} = a_{n-1}\lambda^{n-1}I \Rightarrow B_{n-2} - AB_{n-1} = a_{n-1}I$$

$$B_{n-3}\lambda^{n-2} - AB_{n-2}\lambda^{n-2} = a_{n-2}\lambda^{n-2}I \Rightarrow B_{n-3} - AB_{n-2} = a_{n-2}I$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$B_0\lambda - AB_1\lambda = a_1\lambda I \Rightarrow B_0 - AB_1 = a_1 I$$

$$-B_0A = a_0 I$$

Multiplicando as equações matriciais por $A^n, A^{n-1}, \ldots, A, I$, respectivamente temos:

$$B_{n-1}A^n = A^n$$

$$B_{n-2}A^{n-1} - B_{n-1}A^n = A^{n-1}a_{n-1}$$

$$B_{n-3}A^{n-2} - B_{n-2}A^{n-1} = A^{n-2}a_{n-2}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$B_{n-3}A^{n-2} - B_{n-2}A^{n-1} = A^{n-2}a_{n-2}$$

$$B_0A - B_1A^2 = Aa_1$$

$$-B_0A = a_0I$$

Como temos n equações, a soma dos primeiros termos destas equações será igual a soma dos segundos termos das mesmas:

$$B_{n-1}A^n + B_{n-2}A^{n-1} - B_{n-1}A^n + B_{n-3}A^{n-2} - B_{n-2}A^{n-1} + \ldots + B_0A - B_1A^2 - B_0A =$$

$$= A^n + A^{n-1}a_{n-1} + A^{n-2}a_{n-2} + \ldots + Aa_1 + a_0I$$

$$\Rightarrow A^n + A^{n-1}a_{n-1} + A^{n-2}a_{n-2} + \ldots + Aa_1 + a_0I = 0$$

Temos que

$$\Delta(A) = A^n + A^{n-1}a_{n-1} + A^{n-2}a_{n-2} + \ldots + Aa_1 + a_0I = 0$$

$$\Rightarrow \Delta(A) = 0$$

Isso completa a demonstração do teorema.

Nós denotamos por $\lambda_1, \lambda_2, \ldots, \lambda_n$ todas as raízes do polinômio característico $P(\lambda)$. Então

$$P(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_n)$$

Teorema 10 *Se $\lambda_1, \lambda_2, \ldots, \lambda_n$ são autovalores da matriz A e $g(\mu)$ é um polinômio escalar, então $g(\lambda_1), g(\lambda_2), \ldots, g(\lambda_n)$ são autovalores de $g(A)$.*

Demonstração:

Sejam $g(\mu)$ um polinômio escalar arbitrário e $\lambda_1, \lambda_2, \ldots, \lambda_n$ todos os autovalores da matriz A. Então

$$g(\mu) = a_0(\mu - \mu_1)(\mu - \mu_2)\ldots(\mu - \mu_l).$$

Queremos encontrar os autovalores de $g(A)$.

$$g(A) = a_0(A - \mu_1 I)(A - \mu_2 I)\dots(A - \mu_l I)$$

$$\begin{aligned}\det(g(A)) = \det(a_0 I)\det(A - \mu_1 I)\det(A - \mu_2 I)\dots\det(A - \mu_l I) =\\ a_0^n \det(A - \mu_1 I)\det(A - \mu_2 I)\dots\det(A - \mu_l I) =\\ a_0^n(-1)^{ln}\det(\mu_1 I - A)\det(\mu_2 I - A)\dots\det(\mu_l I - A) =\\ a_0^n(-1)^{ln}\Delta(\mu_1)\Delta(\mu_2)\dots\Delta(\mu_l) =\\ a_0^n(-1)^{ln}\prod_{i=1}^{l}\prod_{k=1}^{n}(\mu_i - \lambda_k) =\\ g(\lambda_1)g(\lambda_2)\dots g(\lambda_n)\end{aligned}$$

Tomamos a equação $\det(g(A)) = g(\lambda_1)g(\lambda_2)\dots g(\lambda_n)$ e substituímos $g(\mu)$ por $\lambda - g(\mu)$. Temos que $\det(\lambda I - g(A)) = [\lambda - g(\lambda_1)][\lambda - g(\lambda_2)]\cdots[\lambda - g(\lambda_n)]$, ou seja, $g(\lambda_i)$ são os autovalores de $g(A)$.

Isso completa a demonstração.

Em particular, se A tem autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ então A^k tem como autovalores $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k (k = 0, 1, 2, \dots)$.

Para encontrarmos os autovetores associados aos autovalores (λ_i) de A precisamos resolver a equação $Av = \lambda v$, ou seja, resolver o sistema:

$$A.\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \lambda_i \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \quad \textit{com} \;\; v = (v_1, v_2, \dots v_n)$$

Definição 25 *Um polinômio escalar $f(\lambda)$ é chamado de polinômio anulador da matriz quadrada A se*

$$f(A) = 0.$$

Um polinômio anulador $\psi(\lambda)$ de grau mínimo com coeficiente do termo de maior grau igual a 1 é chamado de polinômio minimal de A

Pelo teorema de Hamilton-Cayley o polinômio característico $\Delta(\lambda)$ é um polinômio anulador de A. Entretanto, nem sempre ele é um polinômio minimal.

Teorema 11 *Todo polinômio anulador de uma matriz é divisível pelo polinômio minimal.*

Demonstração:

Sejam $f(\lambda)$ e $\psi(\lambda)$ os polinômios anulador e minimal, respectivamente, da matriz A. Dividindo $f(\lambda)$ por $\psi(\lambda)$ temos

$$f(\lambda) = \psi(\lambda)q(\lambda) + r(\lambda),$$

onde r é o resto da divisão, logo tem grau menor que ψ. Substituindo λ pela matriz A temos

$$f(A) = \psi(A)q(A) + r(A).$$

Como $f(A) = 0$ e $\psi(A) = 0$ temos que $r(A) = 0$. Mas o grau de $r(\lambda)$ é menor que o grau do polinômio minimal $\psi(\lambda)$. Então $r(\lambda) = 0$.

Isso prova o teorema.

Iremos agora encontrar a relação entre polinômio minimal e polinômio característico:

Seja $B(\lambda)$ a matriz adjunta de $(\lambda I - A)$, ou seja,

$$(\lambda I - A)B(\lambda) = \Delta(\lambda)I$$

Denotamos por $D_{n-1}(\lambda)$ o máximo divisor comum de todos os elementos de $B(\lambda)$. Então podemos escrever

$$B(\lambda) = D_{n-1}(\lambda)C(\lambda),$$

onde $C(\lambda)$ é chamada de matriz adjunta reduzida de $\lambda I - A$. Isso implica que

$$\Delta(\lambda)I = (\lambda I - A)D_{n-1}(\lambda)C(\lambda).$$

Como $D_{n-1}(\lambda)$ é um polinômio, podemos escrever a equação acima como:

$$\Delta(\lambda)I = (\lambda I - A)C(\lambda)D_{n-1}(\lambda),$$

ou ainda,

$$\frac{\Delta(\lambda)I}{D_{n-1}(\lambda)} = (\lambda I - A)C(\lambda).$$

Chamamos $\frac{\Delta(\lambda)}{D_{n-1}(\lambda)}$ de $\psi(\lambda)$, ou seja, $\frac{\Delta(\lambda)}{D_{n-1}(\lambda)} = \psi(\lambda)$, onde $\psi(\lambda)$ é um polinômio. Segue que:

$$\psi(\lambda)I = \frac{\Delta(\lambda)}{D_{n-1}(\lambda)}I = (\lambda I - A)C(\lambda).$$

Como $\psi(\lambda)I$ é divisível pela esquerda, sem resto, por $(\lambda I - A)$ temos, pelo Teorema de Bézout generalizado, que

$$\psi(A) = \mathbf{0}.$$

Logo, o polinômio $\psi(\lambda) = \frac{\Delta(\lambda)}{D_{n-1}(\lambda)}$ é um polinômio anulador de A. Precisamos agora mostrar que $\psi(\lambda)$ é o polinômio minimal de A.

Denotamos por $\psi^*(\lambda)$ o polinômio minimal. Então $\psi(\lambda)$ é divisível por $\psi^*(\lambda)$:

$$\psi(\lambda) = \psi^*(\lambda)\alpha(\lambda).$$

Como $\psi^*(A) = \mathbf{0}$, a matriz polinomial $\psi^*(\lambda)I$ é divisível pela esquerda por $\lambda I - A$:

$$\psi^*(\lambda)I = (\lambda I - A)C^*(\lambda)$$

$$\psi(\lambda)I = (\lambda I - A)C^*(\lambda)\alpha(\lambda).$$

Temos que, tanto $C(\lambda)$ quanto $C^*(\lambda)$, são quocientes à esquerda na divisão de $\psi(\lambda)$ por $(\lambda I - A)$. Pela unicidade da divisão

$$C(\lambda) = C^*(\lambda)\alpha(\lambda).$$

Segue que $\alpha(\lambda)$, o qual é um divisor comum de todos os elementos da matriz adjunta $C(\lambda)$, é igual a 1, pois a matriz foi obtida através da divisão de $B(\lambda)$ por $D_{n-1}(\lambda)$. Portanto $\alpha(\lambda) = cte$. Como os coeficientes do termos de maior grau de $\psi(\lambda)$ e $\psi^*(\lambda)$ são iguais, concluímos pela equação $\psi(\lambda) = \psi^*(\lambda)\alpha(\lambda)$, temos que $\alpha(\lambda) = 1$, ou seja, $\psi(\lambda) = \psi^*(\lambda)$. Estabelecemos então como relação entre polinômio minimal e polinômio característico:

$$\psi(\lambda) = \Delta(\lambda)D_{n-1}(\lambda).$$

Capítulo 3

Matrizes polinomiais

3.1 Transformações elementares de uma matriz polinomial

Seja $A_j = (a_{ik}^{(j)})(i = 1, 2, \ldots m; k = 1, 2, \ldots, n; j = 0, 1, \ldots, l)$. Podemos representar a matriz polinomial na forma de um polinômio em λ com coeficientes matriciais:

$$A(\lambda) = A_0\lambda^l + A_1\lambda^{l-1} + \ldots + A_{l-1}\lambda + A_l.$$

Podemos fazer as seguintes operações elementares na matriz $A(\lambda)$:

1. Multiplicar uma linha qualquer, por exemplo a i-ésima, por $c \neq 0$, $c \in \mathbb{C}$.
2. Somar uma linha qualquer, por exemplo a i-ésima, a qualquer outra, por exemplo a j-ésima, multiplicada por um polinômio arbitrário $b(\lambda)$.
3. Permutar quaisquer duas linhas, por exemplo a i-ésima e a j-ésima.

Tais operações são equivalentes à multiplicação da matriz $A(\lambda)$, pela esquerda, pelas respectivas matrizes quadradas de ordem m, abaixo:

$$E' = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$E'' = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ b(\lambda) & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$E''' = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix},$$

ou seja, $A(\lambda)$ é transformada em $E'A(\lambda)$, $E''A(\lambda)$ e $E'''A(\lambda)$, respectivamente. Essas operações são chamadas de operações elementares por linha.

Da mesma maneira, podemos definir operações elementares por colunas utilizando colunas ao invés de linhas. O resultado desta aplicação é equivalente a multiplicar a matriz $A(\lambda)$, pela direita, por uma das matrizes E', E'' ou E'''. As matrizes E', E'', E''' são chamadas de matrizes elementares. Essas são matrizes com determinantes constantes. Por isso podemos dizer que o determinante de matrizes elementares não depende de λ e é diferente de zero.

Cada operação elementar por linha (coluna) tem uma operação inversa, a qual é uma operação elementar por linha (coluna).

Se $B(\lambda)$ pode ser obtida de $A(\lambda)$ por meio de operações elementares por linha então podemos escrever

$$B(\lambda) = E_p E_{p-1} \ldots E_1 A(\lambda),$$

ou ainda,

$$B(\lambda) = P(\lambda)A(\lambda),$$

onde $P(\lambda) = E_p E_{p-1} \ldots E_1$ e cada matriz $E_1, E_2, \ldots E_p$ é uma matriz elementar. No caso de equivalência por coluna, podemos escrever

$$B(\lambda) = A(\lambda)Q(\lambda).$$

No caso de equivalência temos

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda),$$

onde $P(\lambda)$ e $Q(\lambda)$ são matrizes polinomiais quadradas com determinantes constantes diferentes de zero.

Segue então a seguinte definição:

Definição 26 *Duas matrizes polinomiais retangulares $A(\lambda)$ e $B(\lambda)$ são chamadas de:*

1. *equivalente por linha se $B(\lambda) = P(\lambda)A(\lambda)$,*
2. *equivalente por coluna se $B(\lambda) = A(\lambda)Q(\lambda)$,*
3. *equivalente se $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$,*

onde $P(\lambda)$ e $Q(\lambda)$ são matrizes quadradas polinomiais com determinantes constantes diferentes de zero.

3.2 Forma canônica de uma matriz polinomial

Seja $A(\lambda)$ uma matriz polinomial retangular. Encontramos a forma canônica de $A(\lambda)$ por meio de operações elementares, ou seja, da mesma forma que encontramos a forma canônica de uma matriz com coeficientes escalares.

Teorema 12 *Uma matriz polinomial retangular arbitrária, de dimensão $m \times n$ pode ser transformada em uma matriz triangular superior*

$$\begin{pmatrix} b_{11}(\lambda) & b_{12}(\lambda) & \cdots & b_{1m}(\lambda) & \cdots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2m}(\lambda) & \cdots & b_{2n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_{mm}(\lambda) & \cdots & b_{mn}(\lambda) \end{pmatrix} (m \leq n)$$

ou

$$\begin{pmatrix} b_{11}(\lambda) & b_{12}(\lambda) & \cdots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_{nn}(\lambda) \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} (m \geq n)$$

por meio de operações elementares por linha, onde os polinômios $b_{1k}(\lambda), b_{2k}(\lambda), \ldots, b_{k-1,k}(\lambda)$ possuem graus menores que o grau de $b_{kk}(\lambda)$, desde que $b_{kk}(\lambda) \not\equiv 0$, e são todos iguais a zero se $b_{kk}(\lambda) = cte \neq 0$ $(k = 2, 3, \ldots, min(m, n))$.

Demonstração:

Assumimos que a primeira coluna de $B(\lambda)$ contém elementos não identicamente iguais a zero. Tomamos o polinômio de menor grau e, através de permutação entre linhas, levamos tal polinômio para elemento $b_{11}(\lambda)$. Então dividimos $b_{i1}(\lambda)$ por $b_{11}(\lambda)$ e obtemos:

$$b_{i1}(\lambda) = b_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda) \quad (i = 1, 2, \ldots, n)$$

onde $q_{i1}(\lambda)$ e $r_{i1}(\lambda)$ são o quociente e o resto, respectivamente. Subtraímos da i-ésima linha a primeira multiplicada por $q_{i1}(\lambda)(i = 1, 2, \ldots, m)$. Caso nem todos os restos $r_{i1}(\lambda)$ sejam identicamente iguais a zero, escolhemos um que não seja igual a zero e seja de menor grau e levamos para o elemento $b_{11}(\lambda)$, utilizando a permutação de linhas. Como resultado destas operações, ou o grau do polinômio $b_{11}(\lambda)$ é reduzido, ou a coluna é zerada.

Repetimos esse processo. Se o grau do polinômio $b_{11}(\lambda)$ for n no início, então ele passa a ser constante ou nulo depois de, no máximo, n passos. Caso $b_{11}(\lambda) = 0$ temos que todos os $b_{ij}(\lambda)$ também são zero. Se $b_{11}(\lambda) = cte \neq 0$, operações elementares por linha anulam o resto da primeira coluna. Portanto a primeira coluna pode ser posta na forma adequada. Para fazermos o mesmo com a segunda coluna, tomamos o elemento $b_{22}(\lambda)$ e aplicamos o mesmo procedimento para as linhas $2, 3, \ldots, m$, até conseguirmos transformar $b_{32}(\lambda), \ldots, b_{m2}(\lambda)$, em elementos de grau zero. Se o polinômio $b_{22}(\lambda)$ não é identicamente igual a zero, por uma aplicação elementar de linha podemos fazer com que o grau de $b_{12}(\lambda)$ se torne menor que o grau de $b_{22}(\lambda)$. Caso o grau de $b_{22}(\lambda)$ seja uma constante não nula, podemos utilizar uma operação por linha para alterar $b_{12}(\lambda)$. Continuando com processo chegaremos à uma das matrizes enunciadas no teorema 12.

Isso completa a demonstração do teorema.

Teorema 13 *Uma matriz polinomial retangular arbitrária, de dimensão $m \times n$, pode ser transformada em uma matriz triangular inferior*

$$\begin{pmatrix} c_{11}(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \cdots & c_{mm}(\lambda) & 0 & \cdots & 0 \end{pmatrix} (m \leq n)$$

ou

$$\begin{pmatrix} c_{11}(\lambda) & 0 & \cdots & 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ c_{n1}(\lambda) & c_{n2}(\lambda) & \cdots & c_{nn}(\lambda) \\ \vdots & \vdots & \vdots & \vdots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \cdots & c_{mm}(\lambda) \end{pmatrix} (m \leq n)$$

por meio de operações elementares por coluna, onde os polinômios $c_{k1}(\lambda), c_{k2}(\lambda), \dots, c_{k,k-1}(\lambda)$ *possuem graus menores que o grau do polinômio* $c_{kk}(\lambda)$, *desde que* $c_{kk}(\lambda) \not\equiv 0$, *e são todos iguais a zero se* $c_{kk}(\lambda) = cte \neq 0$ $(k = 2, 3, \dots, min(m, n))$.

Demonstração:

Essa demonstração é similar à demonstração do teorema anterior. Ao invés de usarmos operações elementares por linhas, faremos agora operações elementares por colunas.

Destes dois teoremas segue o seguinte corolário:

Corolário 13.1 *Se o determinante de uma matriz polinomial quadrada* $P(\lambda)$ *é independente de* λ *e diferente de zero, então a matriz pode ser representada na forma de produto de um número finito de matrizes elementares.*

Demonstração:

Seja $P(\lambda)$ uma matriz polinomial quadrada. Então $P(\lambda)$ pode ser transformada em uma matriz

$$P'(\lambda) = \begin{pmatrix} b_{11}(\lambda) & b_{11}(\lambda) & \cdots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \cdots & b_{2n}(\lambda) \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & b_{nn}(\lambda) \end{pmatrix}$$

por meio de operações elementares por linha, isto é, $P'(\lambda) = E_s \dots E_2 E_1 P(\lambda)$. Seja C o determinante da matriz $P(\lambda)$.Temos que $|C| = |\det P(\lambda)| = |\det P'(\lambda)|$, pois $|\det E_i| = 1 (i = 1, \dots, s)$. Como $P'(\lambda)$ é uma matriz triangular, seu determinante é o produto dos elementos da diagonal. Mas este produto é um produto de polinômios em λ e cujo valor não depende de λ. Logo este produto tem grau 0, ou seja, todos os elementos da diagonal são números. Então, tanto o determinante de $P(\lambda)$ quanto o determinante de $P'(\lambda)$ não dependem de λ e são diferentes de zero. Podemos então, transformar $P'(\lambda)$ em I por meio de operações elementares por linha, ou seja,

$$I = E_{s+L} \dots E_{s+1} P'(\lambda).$$

Temos que:

$$I = E_{s+L} \dots E_{s+1} E_s \dots E_2 E_1 P(\lambda).$$

Como a inversa de uma matriz elementar é uma matriz elementar, temos:

$$P(\lambda) = I(E_1)^{-1}(E_2)^{-1} \dots (E_s)^{-1}(E_{s+1})^{-1} \dots (E_{s+L})^{-1}.$$

Isso completa a demonstração.

Definição 27 *Uma matriz polinomial retangular é chamada de matriz diagonal canônica se é da forma*

$$\begin{pmatrix} a_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & a_2(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_s(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}, s \leq min(m, n)$$

onde:

1. *os polinômios $a_1(\lambda), a_2(\lambda), \ldots, a_s(\lambda)$ não são identicamente iguais a zero,*

2. *cada um dos polinômios $a_2(\lambda), \ldots, a_s(\lambda)$ é divisível por seu antecessor.*

3. *o coeficiente do termo de maior grau de todos os polinômios $a_1(\lambda), a_2(\lambda), \ldots, a_s(\lambda)$ é igual a* 1.

Teorema 14 *Uma matriz polinomial retangular arbitrária $A(\lambda)$ é equivalente a uma matriz diagonal canônica.*

Demonstração:

Escolhemos, entre os elementos $a_{ik}(\lambda)$ de $A(\lambda)$ que não são identicamente iguais a zero, aquele que possue o menor grau em λ e, por permutação de linhas e colunas, podemos levá-lo à posição do elemento $a_{11}(\lambda)$. Encontramos os quocientes e os restos dos polinômios $a_{i1}(\lambda)$ e $a_{ik}(\lambda)$ na divisão por $a_{11}(\lambda)$:

$$a_{i1}(\lambda) = a_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda), \;\; a_{1k}(\lambda) = a_{11}(\lambda)q_{1k}(\lambda) + r_{1k}(\lambda)$$

$$(i = 2, 3, \ldots, n).$$

Se o menor dos restos $r_{i1}(\lambda)$, $r_{1k}(\lambda)$ $(i = 2, 3, \ldots, m;\; k = 2, 3, \ldots, n)$, por exemplo $r_{1k}(\lambda)$, não for identicamente igual a zero, então subtraindo da k-ésima coluna, a primeira coluna multiplicada por $q_{1k}(\lambda)$, substituímos $a_{1k}(\lambda)$ pelo resto $r_{1k}(\lambda)$, o qual possui grau menor que o grau de $a_{11}(\lambda)$. Podemos reduzir o grau do elemento no canto superior esquerdo da matriz colocando em seu lugar um elemento de grau menor em λ.

Mas se todos os restos $r_{21}(\lambda), \ldots, r_{1n}(\lambda)$ são identicamente iguais a zero, subtraíndo da i-ésima linha, a primeira linha multiplicada por $q_{i1}(\lambda)$ $(i = 1, 2, \ldots, m)$ e da k-ésima coluna, a primeira multiplicada por $q_{1k}(\lambda)$ $(k = 2, \ldots, n)$, nós reduzimos a matriz polinomial à forma

$$\begin{pmatrix} a_{11}(\lambda) & 0 & \ldots & 0 \\ 0 & a_{22}(\lambda) & \ldots & a_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & a_{m2}(\lambda) & \ldots & a_{mn}(\lambda) \end{pmatrix}.$$

Se o menor dos elementos $a_{ik}(\lambda)$ $(i = 2, \ldots, m;\ k = 2, \ldots, n)$ não é divisível sem resto por $a_{11}(\lambda)$, então pela adição de primeira coluna à coluna que contém o elemento obtido acima, podemos colocar no lugar do elemento $a_{11}(\lambda)$ um polinômio de grau menor.

Como o elemento original $a_{11}(\lambda)$ tem um grau definido e o processo de redução deste grau não pode ser infinito, podemos, depois de um número finito de operações elementares, obter uma matriz da forma

$$\begin{pmatrix} a_{1}(\lambda) & 0 & \ldots & 0 \\ 0 & b_{22}(\lambda) & \ldots & b_{2n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots \\ 0 & b_{m2}(\lambda) & \ldots & b_{mn}(\lambda) \end{pmatrix},$$

na qual todos os elementos $b_{ik}(\lambda)$ são divisíveis sem resto por $a_1(\lambda)$. Se entre estes elementos $b_{ik}(\lambda)$ há um não indenticamente igual a zero, continuamos o processo de redução com as linhas $2, \ldots, m$ e as colunas $2, \ldots, n$. Conseguimos então, reduzir a matriz acima à uma matriz da forma

$$\begin{pmatrix} a_{1}(\lambda) & 0 & 0 & \ldots & 0 \\ 0 & a_{2}(\lambda) & 0 & \ldots & 0 \\ 0 & 0 & c_{33}(\lambda) & \ldots & c_{3n}(\lambda) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & c_{m3}(\lambda) & \ldots & c_{mn}(\lambda) \end{pmatrix},$$

onde $a_2(\lambda)$ é divisível sem resto por $a_1(\lambda)$ e todos os polinômios $c_{ik}(\lambda)$ são divisíveis sem resto por $a_2(\lambda)$. Continuando o processo chegamos a uma matriz da forma

$$\begin{pmatrix} a_{1}(\lambda) & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & a_{2}(\lambda) & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & a_s(\lambda) & 0 & \ldots & 0 \end{pmatrix},$$

onde os polinômios $a_1(\lambda)$, $a_2(\lambda)$,...,$a_s(\lambda)$ $s \leq min(m,n)$) não são identicamente iguais a zero e cada um é divisível por seu antecessor.

Multiplicando as primeiras s linhas por fatores numéricos diferentes de zero, percebemos que os coeficientes dos termos de maior grau dos polinômios $a_1(\lambda)$, $a_2(\lambda)$,...,$a_s(\lambda)$ são iguais a 1.

Isso completa a demonstração.

3.3 Polinômios invariantes e divisores elementares de uma matriz polinomial

Em $\mathbb{C}$ podemos escrever um polinômio $P(\lambda)$ como $P(\lambda) = \alpha_0(\lambda - \alpha_1)\dots(\lambda - \alpha_k)$. Dizemos que $P(\lambda)$ divide um polinômio $Q(\lambda)$ se e somente se todas as raízes de $P(\lambda)$ são raízes de $Q(\lambda)$. Então o máximo divisor comum entre $P_1, \dots, P_n$ é o polinômio $(\lambda - \beta_1)(\lambda - \beta_2)\dots(\lambda - \beta_j)$, onde $\beta_1, \dots, \beta_j$ são as raízes comuns destes polinômios.

Seja $A(\lambda)$ uma matriz polinomial de rank r, isto é, a matriz tem cofatores de ordem r não identicamente iguais a zero, mas todos os cofatores de ordem maior que r são iguais a zero em λ.

Chamamos de $D_j(\lambda)$ $(j = 1, 2, \dots, r)$ o maior divisor comum de todos os cofatores de ordem j em $A(\lambda)$. Então $D_j(\lambda) = 1\lambda^R + a_{R-1}\lambda^{R-1} + \dots + a_1\lambda + a_0$ e, na série

$$D_r(\lambda), D_{r-1}(\lambda), \dots, D_1(\lambda), D_0(\lambda) \equiv 1$$

cada polinômio é divisível por seu antecessor. Denotamos os quocientes da seguinte forma:

$$i_1 = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, i_2 = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \dots, i_r = \frac{D_1(\lambda)}{D_0(\lambda)} = D_1(\lambda)$$

Definição 28 *Os polinômios $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ definidos acima são chamados de polinômios invariantes de uma matriz retangular $A(\lambda)$.*

Teorema 15 *A matriz polinomial retangular $A(\lambda)$ é equivalente a matriz diagonal canônica*

$$\begin{pmatrix} i_r(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & i_{r-1}(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & i_1(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix},$$

onde $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$ *são polinômios invariantes de* $A(\lambda)$. *Além disto,* r *é o rank de* $A(\lambda)$.

Demonstração:

Sejam $A(\lambda)$ e $B(\lambda)$ duas matrizes polinomiais equivalentes de ordem $m \times n$, ou seja, existem $P(\lambda)$ e $Q(\lambda)$ tais que

$$B(\lambda) = P(\lambda)A(\lambda)Q(\lambda).$$

Sabemos ainda que, em relação a $A(\lambda)$ e $B(\lambda)$,uma pode ser obtida a partir da outra por meio de operações elementares. Aplicando tais operações à $A(\lambda)$, seu rank e seus cofatores, $D_1(\lambda), D_2(\lambda), \ldots, D_r(\lambda)$ permanecem inalterados. Como conseqüência da fórmula de Binet-Cauchy temos que o rank do produto de matrizes triangulares não excede o rank dos fatores. Então

$$r_B \leq min(r_P, r_A, r_Q),$$

onde r_X é o rank da matriz $X(\lambda)$. Logo, todos os cofatores de ordem maior que r_B são iguais a zero. Denotamos por $D^*(\lambda)$ o maior divisor comum de todos os cofatores de $B(\lambda)$. $D^*(\lambda)$ é divisível por $D_p(\lambda)$, $(p = 1, 2, \ldots, min(m, n))$, onde $D_p(\lambda)$ é cofator de $A(\lambda)$.

Mas também podemos obter $B(\lambda)$ a partir de $A(\lambda)$. Então o rank de $B(\lambda)$ e seus cofatores, $D_1^*(\lambda), D_2^*(\lambda), \ldots, D_r^*(\lambda)$ permanecem inalterados. Chamamos de $D(\lambda)$ máximo divisor comum de todos os cofatores de $A(\lambda)$. Logo $D(\lambda)$ é divisível por $D_p^*(\lambda)$, $(p = 1, 2, \ldots, min(m, n))$, onde $D_p^*(\lambda)$ é cofator de $B(\lambda)$. Logo,

$$r_A = r_B; D_1(\lambda) = D_1^*(\lambda), \ldots, D_r(\lambda) = D_r^*(\lambda).$$

Como os polinômios $D_1(\lambda), D_2(\lambda), \ldots, D_r(\lambda)$ permanecem inalterados, podemos afirmar que $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$ definidos anteriormente também permanecem invariantes na transformação de uma matriz em outra equivalente a ela. Quando a matriz polinomial tem a forma diagonal canônica

$$\begin{pmatrix} a_1(\lambda) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & a_2(\lambda) & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_s(\lambda) & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}, s \leq min(m, n)$$

então para essa matriz temos:

$$D_1(\lambda) = a_1(\lambda), D_2(\lambda) = a_1(\lambda)a_2(\lambda), \ldots, D_r(\lambda) = a_1(\lambda)a_2(\lambda)\ldots a_r(\lambda)$$

Logo,

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)} = \frac{a_1(\lambda)a_2(\lambda)\ldots a_{r-1}(\lambda)a_r(\lambda)}{a_1(\lambda)a_2(\lambda)\ldots a_{r-1}(\lambda)} = a_r(\lambda)$$

$$i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)} = \frac{a_1(\lambda)a_2(\lambda)\ldots a_{r-2}(\lambda)a_{r-1}(\lambda)}{a_1(\lambda)a_2(\lambda)\ldots a_{r-2}(\lambda)} = a_{r-1}(\lambda)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$i_r(\lambda) = D_1(\lambda) = a_1(\lambda)$$

Deste modo $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$ são polinômios invariantes da matriz diagonal, bem como da matriz original $A(\lambda)$ pois esta é equivalente à matriz acima.

Isso completa a demonstração.

Corolário 15.1 *Duas matrizes retangulares de mesma dimensão $A(\lambda)$ e $B(\lambda)$ são equivalentes se e somente se elas têm os mesmos polinômios invariantes.*

Demonstração:

A necessidade desta condição foi explicada acima. A suficiência segue do fato que duas matrizes polinomiais que possuem os mesmos polinômios invariantes são equivalentes a uma única e mesma matriz diagonal canônica.

Corolário 15.2 *Na sequência de polinômios invariantes*

$$i_1 = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, i_2 = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \ldots, i_r = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda) \equiv 1)$$

cada polinômio é divisível por seu antecessor.

Demonstração:

Sejam

$$i_1 = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, i_2 = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \ldots, i_r = \frac{D_1(\lambda)}{D_0(\lambda)} (D_0(\lambda) \equiv 1)$$

os polinômios invariantes da matriz $A(\lambda)$. Mas

$$i_1(\lambda) = a_r(\lambda), i_2(\lambda) = a_{r-1}(\lambda), \ldots, i_r(\lambda) = a_1(\lambda).$$

Pela definição de matriz diagonal canônica, cada polinômio $a_j(\lambda)$ $(j = 1, 2, \ldots, r)$ é divisível por seu antecessor. Logo, cada polinômio invariante divide seu antecessor. Isso completa a demonstração.

Iremos agora indicar um método de calcular os polinômios invariantes de uma matriz polinomial quase-diagonal se os polinômios invariantes das matrizes nos blocos diagonais forem conhecidos.

Introduzimos duas novas notações usadas na demonstração a seguir. Denotaremos uma matriz diagonal, $n \times n$, com elementos a_{11}, a_{22}, ..., a_{nn} na diagonal, por $\{a_{11}, a_{22}, \ldots, a_{nn}\}$. Usaremos $A \sim B$ para dizer que A e B são equivalentes.

Teorema 16 *Se numa matriz retangular quase-diagonal*

$$C(\lambda) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$

todo polinômio invariante de $A(\lambda)$ divide todo polinômio invariante de $B(\lambda)$, então o conjunto dos polinômios invariantes de $C(\lambda)$ é a união dos polinômios invariantes de $A(\lambda)$ e $B(\lambda)$.

Demonstração:

Sejam $i'_1(\lambda), i'_2(\lambda), \ldots, i'_r(\lambda), i''_1(\lambda), i''_2(\lambda), \ldots, i''_q(\lambda)$ os polinômios invariantes das matrizes $A(\lambda)$ e $B(\lambda)$ respectivamente. Então

$$A(\lambda) \sim \{i'_r(\lambda), \ldots, i'_1(\lambda), 0, \ldots, 0\}, B(\lambda) \sim \{i''_q(\lambda), \ldots, i''_1(\lambda), 0, \ldots, 0\}$$

e daí

$$C(\lambda) \sim \{i'_r(\lambda), \ldots, i'_1(\lambda), i''_q(\lambda), \ldots, i''_1(\lambda), 0, \ldots, 0\}.$$

O lado direito desta relação é a forma diagonal canônica da matriz polinomial. Pelo teorema 16 os elementos da diagonal desta matriz que não são identicamente iguais a zero formam um sistema completo de invariantes da matriz polinomial $C(\lambda)$. Isso prova o teorema.

Para determinar os polinômios invariantes de $C(\lambda)$ no caso geral de polinômios arbitrários invariantes de $A(\lambda)$ e $B(\lambda)$ nós usamos o conceito importante de divisores elementares.

Decompomos os polinômios invariantes $i_1(\lambda), i_2(\lambda), \ldots, i_r(\lambda)$ em fatores lineares:

$$i_1(\lambda) = (\lambda - \varphi_1)^{c_{11}}(\lambda - \varphi_2)^{c_{12}} \dots (\lambda - \varphi_s)^{c_{1s}},$$
$$i_2(\lambda) = (\lambda - \varphi_1)^{c_{21}}(\lambda - \varphi_2)^{c_{22}} \dots (\lambda - \varphi_s)^{c_{2s}},$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$i_r(\lambda) = (\lambda - \varphi_1)^{c_{r1}}(\lambda - \varphi_2)^{c_{r2}} \dots (\lambda - \varphi_s)^{c_{rs}}.$$
$$(c_{ij} \geq c_{kj} \quad i \geq k \quad i, k = 1, 2, \dots, r \quad j = 1, 2, \dots, s)$$

$\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_s(\lambda)$ são fatores lineares e distintos de $\mathbb{C}$ (e com maior coeficiente igual a 1) que aparecem em $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$.

Definição 29 *Todas as potências dentre $(\lambda - \varphi_1)^{c_{11}}, \dots, (\lambda - \varphi_s)^{c_{rs}}$, definidas acima, diferentes de 1, são chamadas de divisores elementares da matriz $A(\lambda)$ no conjunto $\mathbb{C}$.*

Teorema 17 *O conjunto dos divisores elementares de uma matriz retangular quase-diagonal*
$$C(\lambda) = \begin{pmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{pmatrix}$$
é obtido pela combinação dos divisores elementares de $A(\lambda)$ com os de $B(\lambda)$.

Demonstração:

Decompomos os polinômios invariantes de $A(\lambda)$ e $B(\lambda)$ em fatores lineares do conjunto $\mathbb{C}$:

$$i'_1(\lambda) = [\lambda - \varphi_1]^{c'_{11}}[\lambda - \varphi_2]^{c'_{12}} \dots [\lambda - \varphi_s]^{c'_{1s}}; i''_1(\lambda) = [\lambda - \varphi_1]^{c''_{11}}[\lambda - \varphi_2]^{c''_{12}} \dots [\lambda - \varphi_s]^{c''_{1s}}$$

$$i'_2(\lambda) = [\lambda - \varphi_1]^{c'_{21}}[\lambda - \varphi_2]^{c'_{22}} \dots [\lambda - \varphi_s]^{c'_{2s}}; i''_2(\lambda) = [\lambda - \varphi_1]^{c''_{21}}[\lambda - \varphi_2]^{c''_{22}} \dots [\lambda - \varphi_s]^{c''_{2s}}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$i'_r(\lambda) = [\lambda - \varphi_1]^{c'_{r1}}[\lambda - \varphi_2]^{c'_{r2}} \dots [\lambda - \varphi_s]^{c'_{rs}}; i''_q(\lambda) = [\lambda - \varphi_1]^{c''_{q1}}[\lambda - \varphi_2]^{c''_{q2}} \dots [\lambda - \varphi_s]^{c''_{qs}}$$

Para cada k, denotamos por c_{ik}, $(i = 1, \dots, m_k)$, todos os números dentre $c'_{1k}, c'_{2k}, \dots, c'_{rk}, c''_{1k}, c''_{2k}, \dots, c''_{qk}$ diferentes de zero, em ordem decrescente. Pelo teorema 15, $C(\lambda)$ é equivalente à matriz $\{i'_r(\lambda), \dots, i'_1(\lambda), \dots, i''_q(\lambda), \dots, i''_1(\lambda), 0, \dots, 0\}$. Portanto, por permutação de linhas e colunas podemos transformar essa matriz em uma matriz diagonal da forma

$$\{[\lambda - \varphi_1]^{c_{11}}.(*), [\lambda - \varphi_1]^{c_{21}}.(*), \dots, [\lambda - \varphi_1]^{c_{m1}}.(*), (**), \dots, (**)\},$$

onde $(*)$ são polinômios primos a $(\lambda - \varphi_1)$ e $(**)$ são polinômios primos a $(\lambda - \varphi_1)$ ou identicamente iguais a zero. Temos então a seguinte decomposição para os polinômios $D_r(\lambda), D_{r-1}(\lambda), \ldots$ e $i_1(\lambda), i_2(\lambda), \ldots$ da matriz $C(\lambda)$:

$$D_r(\lambda) = [\lambda - \varphi_1]^{c_{11}+c_{21}+\ldots+c_{r1}}.(*), D_{r-1}(\lambda) = [\lambda - \varphi_1]^{c_{21}+\ldots+c_{r1}}.(*), \ldots,$$

$$i_1 = [\lambda - \varphi_1]^{c_{11}}(*), i_2 = [\lambda - \varphi_1]^{c_{21}}(*), \ldots$$

Segue que $[\lambda - \varphi_1]^{c_{11}}, [\lambda - \varphi_1]^{c_{21}}, \ldots, [\lambda - \varphi_1]^{c_{r1}}$, ou seja, todas as potências

$$[\lambda - \varphi_1]^{c'_{11}}, \ldots, [\lambda - \varphi_1]^{c'_{r1}}, [\lambda - \varphi_1]^{c''_{11}}, \ldots, [\lambda - \varphi_1]^{c''_{g1}},$$

as quais são diferentes de 1, são divisores elementares de $C(\lambda)$.

Os divisores elementares de $C(\lambda)$, os quais são potências de $[\lambda - \varphi_2]$, são determinados da mesma forma.

Isso completa a demonstração do teorema.

Supomos agora que $A = (a_{ik})_1^n$ é uma matriz com elementos em $\mathbb{C}$. Como a matriz característica de A é dada por

$$(\lambda I - A) = \begin{pmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{pmatrix},$$

onde esta é polinomial de rank n. Seus polinômios invariantes

$$i_1 = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}, i_2 = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \ldots, i_n = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda) \equiv 1)$$

são chamados de polinômios invariantes da matriz A e os divisores elementares correspondentes em $\mathbb{C}$ são chamados de divisores elementares da matriz A em $\mathbb{C}$.

3.4 Equivalência de binômios lineares

Consideremos duas matrizes polinomiais quadradas, $A(\lambda)$ e $B(\lambda)$, de ordem n nas quais os elementos não possuem graus maiores que 1 em λ. Essas matrizes polinomiais podem ser representadas na forma de binômios matriciais:

$$A(\lambda) = A_0\lambda + A_1, \quad B(\lambda) = B_0\lambda + B_1.$$

Vamos assumir que estes binômios são de grau 1 e regulares, isto é, $\det(A_0) \neq 0$ e $\det(B_0) \neq 0$.

Teorema 18 *Se dois binômios regulares $A_0\lambda + A_1$ e $B_0\lambda + B_1$ são equivalentes, então eles são estritamente equivalentes, isto é, na identidade*

$$B_0\lambda + B_1 = P(\lambda)(A_0\lambda + A_1)Q(\lambda)$$

as matrizes $P(\lambda)$ e $Q(\lambda)$, com determinantes constantes diferentes de zero, podem ser substituídas por matrizes constantes não-singulares P e Q:

$$B_0\lambda + B_1 = P(A_0\lambda + A_1)Q$$

Esta identidade é equivalente a duas equações matriciais: $B_0 = PA_0Q$ e $B_1 = PA_1Q$.

Demonstração:

Seja $\det(P(\lambda))$ independente de λ e diferente de zero. Então a matriz inversa $M(\lambda) = P^{-1}(\lambda)$ também é uma matriz polinomial. Podemos escrever:

$$M(\lambda)(B_0\lambda + B_1) = (A_0\lambda + A_1)Q(\lambda)$$

onde $M(\lambda)$ e $Q(\lambda)$ são matrizes polinomiais. Vamos dividir $M(\lambda)$ pela esquerda por $A_0\lambda + A_1$ e $Q(\lambda)$ pela direita por $B_0\lambda + B_1$. Então temos:

$$M(\lambda) = (A_0\lambda + A_1)S(\lambda) + M$$

e

$$Q(\lambda) = T(\lambda)(B_0\lambda + B_1) + Q,$$

onde M e Q são matrizes quandradas constantes (independentes de λ) de ordem n. Daí vem:

$$M(\lambda)(B_0\lambda + B_1) = (A_0\lambda + A_1)Q(\lambda)$$

$$[(A_0\lambda + A_1)S(\lambda) + M](B_0\lambda + B_1) = (A_0\lambda + A_1)[T(\lambda)(B_0\lambda + B_1) + Q]$$

$$(A_0\lambda + A_1)S(\lambda)(B_0\lambda + B_1) + M(B_0\lambda + B_1) = (A_0\lambda + A_1)T(\lambda)(B_0\lambda + B_1) + (A_0\lambda + A_1)Q$$

$$(A_0\lambda + A_1)S(\lambda)(B_0\lambda + B_1) - (A_0\lambda + A_1)T(\lambda)(B_0\lambda + B_1) = (A_0\lambda + A_1)Q - M(B_0\lambda + B_1)$$

$$((A_0\lambda + A_1)[T(\lambda) - S(\lambda)](B_0\lambda + B_1) = M(B_0\lambda + B_1) - (A_0\lambda + A_1)Q.$$

O lado direito da equação acima é de grau 1, pois M e Q são matrizes constantes. Vamos analisar o lado esquerdo da equação. No caso de termos $T(\lambda) - S(\lambda) \neq 0$, teremos então uma equação de grau maior ou igual a 2. Entretanto, isso é impossível devido ao fato de termos uma equação de grau 1 do outro lado. Logo,

$$T(\lambda) - S(\lambda) = 0 \Rightarrow T(\lambda) = S(\lambda).$$

Então temos

$$M(B_0\lambda + B_1) - (A_0\lambda + A_1) = 0,$$

ou seja,

$$M(B_0\lambda + B_1) = (A_0\lambda + A_1)Q.$$

Precisamos mostrar agora que M é uma matriz não-singular. Dividimos $P(\lambda)$, pela esquerda, por $(B_0\lambda + B_1)$. Daí

$$P(\lambda) = (B_0\lambda + B_1)U(\lambda) + P.$$

Então temos:

$$I = M(\lambda)P(\lambda) = M(\lambda)(B_0\lambda + B_1)U(\lambda) + M(\lambda)P =$$

$$= (A_0\lambda + A_1)Q(\lambda)U(\lambda) + M(\lambda)P = (A_0\lambda + A_1)Q(\lambda)U(\lambda) + (A_0\lambda + A_1)S(\lambda)P + MP =$$

$$= (A_0\lambda + A_1)[Q(\lambda)U(\lambda) + S(\lambda)P] + MP,$$

ou seja,

$$I = (A_0\lambda + A_1)[Q(\lambda)U(\lambda) + S(\lambda)P] + MP.$$

Esta equação é de grau zero em λ (pois é igual a I). Então a expressão entre colchetes é identicamente igual a zero. Logo temos:

$$I = (A_0\lambda + A_1).0 + MP \Rightarrow$$

$$MP = I$$

onde $\det(M) \neq 0$ e $M^{-1} = P$. Multiplicando ambos os lados da equação

$$M(B_0\lambda + B_1) = (A_0\lambda + A_1)Q$$

pela esquerda por P obtemos:

$$PM(B_0\lambda + B_1) = P(A_0\lambda + A_1)Q$$

$$PP^{-1}(B_0\lambda + B_1) = P(A_0\lambda + A_1)Q$$

$$(B_0\lambda + B_1) = P(A_0\lambda + A_1)Q.$$

O fato de P ser não-singular se deve a $MP = I$. O fato de P e Q serem não singulares vem de $(B_0\lambda + B_1) = P(A_0\lambda + A_1)Q$. Então a identidade acima implica que:

$$B_0 = PA_0Q$$

e daí vem:

$$\det(P)\det(A_0)\det(Q) = \det(B_0) \neq 0.$$

Isso completa a demonstração deste teorema.

Nota: Pela demonstração deste teorema temos que as matrizes constantes P e Q, que substituem as matrizes polinomiais $P(\lambda)$ e $Q(\lambda)$, são restos, esquerdo e direito, respectivamente, de $P(\lambda)$ e $Q(\lambda)$ na divisão por $B_0(\lambda) + B_1$.

Capítulo 4

Matrizes semelhantes

4.1 Um critério para a semelhança entre matrizes

Seja $A = (a_{ik})_1^n$ uma matriz com elementos numéricos de $\mathbb{C}$. Sua matriz característica $(\lambda I - A)$ é uma matriz polinomial de rank n e tem polinômios invariantes:

$$i_1(\lambda), i_2(\lambda), \ldots, i_n(\lambda).$$

Teorema 19 *Duas matrizes $A = (a_{ik})_1^n$ e $B = (b_{ik})_1^n$ são semelhantes ($B = T^{-1}AT$) se e somente se elas possuem os mesmos polinômios invariantes ou, o que significa o mesmo, os mesmos divisores elementares.*

Demonstração:

Sejam A e B duas matrizes semelhantes. Então existe uma matriz não-singular T tal que

$$B = T^{-1}AT.$$

Temos que

$$\lambda I - B = T^{-1}(\lambda I - A)T.$$

Esta equação mostra que as matrizes características são equivalentes. Então, pelo corolário 15.1, A e B têm os mesmos polinômios invariantes. Pela definição de divisores elementares, concluímos que A e B possuem os mesmos divisores elementares.

Sejam A e B duas matrizes polinomiais. Supomos que as matrizes características $(\lambda I - A)$ e $(\lambda I - B)$ têm os mesmos polinômios invariantes. Então, pelo corolário 15.1, elas são equivalentes, ou seja, existem duas matrizes polinomiais $P(\lambda)$ e $Q(\lambda)$ tais que:

$$\lambda I - B = P(\lambda)(\lambda I - A)Q(\lambda).$$

Aplicando o teorema 18 na identidade acima, podemos substituir $P(\lambda)$ e $Q(\lambda)$ por matrizes constantes P e Q:

$$\lambda I - B = P(\lambda I - A)Q.$$

Pelo Teorema de Bézout generalizado temos

$$P = \hat{P}(B) \quad e \quad Q = Q(B).$$

Daí vem:

$$\lambda I - B = P(\lambda I - A)Q$$

$$\lambda I - B = \hat{P}(B)(\lambda I - A)Q(B).$$

Então:

$$B = PAQ \quad e \quad I = PQ,$$

isto é,

$$B = T^{-1}AT,$$

onde $T = Q = P^{-1}$.

Isso completa a demonstração do teorema.

Suplemento do Teorema 19 *Se $A = (a_{ik})_1^n$ e $B = (b_{ik})_1^n$ são duas matrizes semelhantes, ou seja,*

$$B = T^{-1}AT,$$

então nós podemos escolher como matriz transformação T, a matriz

$$T = Q(B) = [\hat{P}(B)]^{-1},$$

onde $P(\lambda)$ e $Q(\lambda)$ são matrizes polinomiais na identidade

$$\lambda I - B = P(\lambda)(\lambda I - A)Q(\lambda)$$

que relacionam as matrizes características equivalentes $\lambda I - A$ e $\lambda I - B$.

4.2 A forma canônica de uma matriz

Seja

$$g(\lambda) = \lambda^m + \alpha_1\lambda^{m-1} + \ldots + \alpha_{m-1}\lambda + \alpha_m$$

um polinômio com coeficientes em $\mathbb{C}$.

Consideramos a matriz quadrada L de ordem m

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_m \\ 1 & 0 & \cdots & 0 & -\alpha_{m-1} \\ 0 & 1 & \cdots & 0 & -\alpha_{m-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & -\alpha_1 \end{pmatrix}.$$

O polinômio característico de L é dado por

$$\det(\lambda I - L) = \begin{vmatrix} \lambda & 0 & \cdots & 0 & \alpha_m \\ -1 & \lambda & \cdots & 0 & \alpha_{m-1} \\ 0 & -1 & \cdots & 0 & \alpha_{m-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -1 & \lambda + \alpha_1 \end{vmatrix} = \lambda^m + \alpha_1 \lambda^{m-1} + \ldots + \alpha_{m-1}\lambda + \alpha_m = g(\lambda).$$

O cofator do elemento α_m no determinante da matriz característica é igual a ± 1. Portanto $D_{m-1}(\lambda) = 1$ e $i_1 = \frac{D_m(\lambda)}{D_{m-1}(\lambda)} = D_m(\lambda) = g(\lambda)$, $i_2(\lambda) = \ldots = i_m(\lambda) = 1$.

Chamamos a matriz L de matriz companheira do polinômio $g(\lambda)$.

Seja $A = (a_{ik})_1^n$ uma matriz cujos polinômios invariantes são

$$i_1(\lambda), i_2(\lambda), \ldots, i_t(\lambda), i_{t+1}(\lambda) = \ldots = i_n(\lambda) = 1,$$

onde os polinômios $i_1(\lambda), i_2(\lambda), \ldots, i_t(\lambda)$ têm graus positivos e cada um divide seu antecessor. Denotamos as matrizes companheiras por $L_1, L_2, \ldots, L_t$.

Então, pelo teorema 17 a matriz quase-diagonal de ordem n,

$$L_I = \{L_1, L_2, \ldots, L_t\},$$

tem os polinômios i_j $(j = 1, 2, \ldots, n)$ como seus polinômios invariantes. Como as matrizes A e L_I têm os mesmos polinômios invariantes, elas são semelhantes, ou seja, sempre existe uma matriz não-singular U tal que

$$A = U L_I U^{-1}.$$

A matriz L_I é chamada de Primeira Forma Canônica Natural da Matriz A. Esta forma canônica é caracterizada por:

1. forma quase-diagonal

$$L_I = \{L_1, L_2, \ldots, L_t\};$$

2. estrutura especial de blocos diagonais, onde tais blocos apresentam a forma da matriz L enunciada acima;

3. na seqüência de polinômios característicos dos blocos diagonais cada polinômio divide seu antecessor.

Denotamos por

$$\chi_1(\lambda), \chi_2(\lambda), \ldots, \chi_u(\lambda)$$

os divisores elementares de $A = (a_{ik})_1^n$. As matrizes companheiras correspondentes serão denotadas por

$$L^{(1)}, L^{(2)}, \ldots, L^{(u)}.$$

Como $\chi_j(\lambda)$ é o único divisor elementar de $L^{(j)}(j = 1, 2, \ldots, u)$, a matriz quase-diagonal

$$L_{II} = \{L^{(1)}, L^{(2)}, \ldots, L^{(u)}\}$$

tem, pelo teorema 17, os polinômios $\chi_1(\lambda), \ldots, \chi_u(\lambda)$ como divisores elementares.

As matrizes A e L_{II} têm os mesmos divisores elementares. Então estas matrizes são semelhantes, isto é, existe uma matriz V, não-singular, tal que

$$A = V L_{II} V^{-1}.$$

A matriz L_{II} é chamada de Segunda Forma Canônica Natural da Matriz A. Esta forma é caracterizada por

1. a forma quase-diagonal L_{II};

2. a estrutura especial de blocos diagonais;

3. o polinômio característico de cada bloco diagonal é uma potência de um polinômio linear.

Nota: Os divisores elementares de uma matriz A, ao contrário dos polinômios invariantes, são essencialmente ligados a um elemento de $\mathbb{C}$.

Supomos, por exemplo, que $A = (a_{ik})_1^n$ é uma matriz com coeficientes reais. Mas este polinômio pode ter raízes complexas. Quando trabalhamos em $\mathbb{R}$, entre os divisores elementares pode haver potências de trinômios quadráticos com coeficientes reais. Quando trabalhamos com o conjunto dos números complexos, então todo divisor elementar tem a forma $(\lambda - \lambda_0)^p$. Consideramos agora não apenas os elementos de A, mas também os seus autovalores. Então os divisores elementares de A têm a forma:

$$(\lambda-\lambda_1)^{p_1},(\lambda-\lambda_2)^{p_2},\ldots,(\lambda-\lambda_u)^{p_u};\quad (p_1+p_2+\ldots+p_u=n).$$

Consideramos um destes divisores elementares,

$$(\lambda-\lambda_0)^{p},$$

e associamos a ele a seguinte matriz de ordem p:

$$\begin{pmatrix} \lambda_0 & 1 & 0 & \cdots & 0 \\ 0 & \lambda_0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & \lambda_0 \end{pmatrix} = \lambda_0 I^{(p)} + H^{(p)},$$

onde I é a matriz identidade e H é uma matriz da forma

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

onde $p=$ ordem de H e $H^{p}=H$.

Essa matriz possui apenas um único divisor elementar, $(\lambda-\lambda_0)^{p}$. A matriz $\lambda_0 I^{(p)}+H^{(p)}$ é chamada de *Bloco de Jordan* correspondente ao divisor elementar $(\lambda-\lambda_0)^{p}$.

Os Blocos de Jordan correspondentes aos divisores elementares $(\lambda-\lambda_i)^{p_i}$ $(i=1,\ldots,u)$ serão denotados por $J_1,J_2,\ldots,J_u$.

Então a matriz quase-diagonal

$$J=\{J_1,J_2,\ldots,J_u\}$$

tem as potências $(\lambda-\lambda_i)^{p_i}$ $(i=1,\ldots,u)$ como seus divisores elementares.

A matriz J pode ser escrita na forma

$$J=\{\lambda_1 I_1+H_1,\lambda_2 I_2+H_2,\ldots,\lambda_u I_u+H_u\};$$

onde $I_k=I^{(p_k)},H_k=H^{(p_k)},\quad (k=1,2,\ldots,u)$.

No caso de as matrizes A e J terem os mesmos divisores elementares elas são semelhantes, isto é, existe uma matriz não-singular T tal que

$$A = TJT^{-1} = T\{\lambda_1 I_1 + H_1, \lambda_2 I_2 + H_2, \ldots, \lambda_u I_u + H_u\}T^{-1}.$$

A matriz J é chamada de Forma Canônica Jordan é caracterizada por sua forma quase-diagonal e por sua estrutura de blocos diagonais. Os esquema abaixo descreve a matriz de Jordan J através de seus divisores elementares $(\lambda-\lambda_1)^2, (\lambda-\lambda_2)^3, (\lambda-\lambda_3)$, $(\lambda-\lambda_4)^2$.

$$J = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 \end{pmatrix}.$$

A Forma de Jordan é uma matriz diagonal se e somente se todos os divisores elementares de uma matriz A são de grau 1, e neste caso temos:

$$A = T\{\lambda_1, \lambda_2, \ldots, \lambda_n\}T^{-1}.$$

No lugar do bloco de Jordan visto anteriormente, algumas vezes usamos o bloco inferior de Jordan de ordem p:

$$\begin{pmatrix} \lambda_0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & \lambda_0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \lambda_0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & \lambda_0 \end{pmatrix} = \lambda_0 I^{(p)} + F^{(p)},$$

onde I é a matriz identidade de ordem p e F é uma matriz, também de ordem p, da forma

$$\begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

A matriz também tem apenas um divisor elementar $(\lambda-\lambda_0)^P$. Para os divisores elementares $(\lambda - \lambda_i)^{P_i} (i = 1, \dots, u)$ existem matrizes inferiores de Jordan correspondentes

$$J_{(1)} = \{\lambda_1 I_1 + F_1, \lambda_2 I_2 + F_2, \dots, \lambda_u I_u + F_u\},$$

$$(I_k = I^{(P_k)}, F_k = F^{(P_k)}; k = 1, 2, \dots, u).$$

Uma matriz arbitrária A que possui como divisores elementares $(\lambda - \lambda_i)^{P_i}$ $(i = 1, \dots, u)$ é semelhante a $J_{(1)}$, isto é, existe uma matriz não-singular T_1 tal que

$$A = T_1 J_{(1)} T_1^{-1} = T_1 \{\lambda_1 I_1 + F_1, \lambda_2 I_2 + F_2, \dots, \lambda_u I_u + F_u\} T_1^{-1}.$$

Se $\lambda_0 \neq 0$ então as matrizes

$$\lambda_0 (I^{(P)} + H^{(P)}) \quad \text{e} \quad \lambda_0 (I^{(P)} + F^{(P)})$$

têm o mesmo divisor elementar $(\lambda - \lambda_0)^{(P_i)}$. Então para uma matriz não-singular A que tenha como divisores elementares $(\lambda - \lambda_i)^{P_i} (i = 1, \dots, u)$ nós temos:

$$A = TJT^{-1} \quad e \quad A = T_1 J_{(1)} T_1^{-1}.$$

Daí concluímos que

$$A = T_2 \{\lambda_1 (I_1 + H_1), \lambda_2 (I_2 + H_2), \dots, \lambda_u (I_u + H_u)\} T_2^{-1},$$

$$A = T_3 \{\lambda_1 (I_1 + F_1), \lambda_2 (I_2 + F_2), \dots, \lambda_u (I_u + F_u)\} T_3^{-1}.$$

4.3 Um método geral de construir uma matriz transformação

Em muitos problemas sobre teoria de matrizes e suas aplicações é suficiente conhecer a forma canônica na qual uma matriz $A = (a_{ik})_1^n$ pode ser conduzida por meio de transformações que envolvem semelhança entre matrizes. A forma canônica é determinada por meio de polinômios invariantes da matriz característica $(\lambda I - A)$. Para encontrá-la utilizaremos fórmulas já definidas ou a redução da matriz característica $\lambda I - A$ para a forma diagonal canônica através de operações elementares. Entretanto, em alguns problemas, é necessário conhecer não apenas a forma canônica $\tilde{A}$ da matriz A, mas também uma matriz transformação não-singular (T).

Um método imediato para encontrarmos T consiste no seguinte:

$$A = T\bar{A}T^{-1}$$

$$AT = T\bar{A}$$

$$AT - T\bar{A} = 0.$$

A equação matricial em T é equivalente a um sistema de n^2 equações lineares homogêneas com n^2 incógnitas em T. A determinação de uma matriz transformação consiste na resolução de um sistema de n^2 equações.

Entretanto precisamos escolher uma solução de modo que $\det(T) \neq 0$. A existência de cada uma das soluções está certa, desde que A e $\bar{A}$ tenham os mesmos polinômios invariantes.

Considerando que a forma canônica é determinada unicamente pela matriz A, para a matriz transformação T nós sempre temos um conjunto enumerável de valores que são dados por

$$T = UT_1,$$

onde T_1 é uma das matrizes transformação e U é uma matriz arbitrária que é permutável com A, ou seja, $UA = AU$. Ou ainda,

$$T = T_1V_1,$$

onde V é uma matriz arbitrária permutável com $\bar{A}$.

Este método é de fácil compreensão, mas é pouco prático devido ao número de equações necessárias para encontrar T.

Partiremos agora para um método mais eficiente de construírmos a matriz transformação T, baseado no suplemento do teorema 19.

Escolhemos como matriz transformação

$$T = Q(\bar{A}).$$

Então temos:

$$\lambda I - \bar{A} = P(\lambda)(\lambda I - A)Q(\lambda).$$

Essa equação expressa a equivalência entre as matrizes características $(\lambda I - A)$ e $(\lambda I - \bar{A})$, e $P(\lambda)$ e $Q(\lambda)$ são matrizes polinomiais com determinantes constantes diferentes de zero, dadas pelo teorema citado acima.

Para encontrarmos $Q(\lambda)$, reduzimos as matrizes $(\lambda I - A)$ e $(\lambda I - \bar{A})$ para a forma canônica por meio de operações elementares correspondentes:

$$\{i_n(\lambda), i_{n-1}(\lambda), \ldots, i_1(\lambda)\} = P_1(\lambda)(\lambda I - A)Q_1(\lambda)$$

$$\{i_n(\lambda), i_{n-1}(\lambda), \ldots, i_1(\lambda)\} = P_2(\lambda)(\lambda I - \tilde{A})Q_2(\lambda)$$

onde $Q_1(\lambda) = T_1T_2\ldots T_{p_1}$, $Q_2 = T_1^*T_2^*\ldots T_{p_2}^*$ e $T_1,\ldots,T_{p_1},T_1^*,\ldots T_{p_2}^*$ são as matrizes elementares correspondentes às operações elementares nas colunas das matrizes polinomias $\lambda I - A$ e $\lambda I - \tilde{A}$. Temos então que

$$\lambda I - \tilde{A} = P(\lambda)(\lambda I - A)Q(\lambda),$$

onde $Q(\lambda) = Q_1(\lambda)Q_2^{-1}(\lambda) = T_1T_2\ldots T_{p_1}T_{p_2}^{*-1}T_{p_2-1}^{*-1}\ldots T_1^{*-1}$.

Podemos encontrar a matriz $Q(\lambda)$ aplicando sucessivamente às colunas da matriz I operações elementares com as matrizes $T_1,\ldots,T_{p_1},T_{p_2}^{*-1},\ldots,T_1^{*-1}$. Depois substituímos o argumento λ em $Q(\lambda)$ pela matriz $\tilde{A}$.

4.4 Outro método para construírmos uma matriz transformação

Este novo método utiliza um número menor de operações que os vistos anteriormente. Entretanto, só podemos utilizá-lo quando a forma canônica de Jordan e os divisores elementares

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \ldots$$

de uma matriz A são conhecidos.

Seja $A = TJT^{-1}$, onde

$$J = \{\lambda_1 I^{p_1} + H^{p_1}, \lambda_2 I^{p_2} + H^{p_2}, \ldots\} = \begin{pmatrix} \lambda_1 & 1 & \cdots & 0 & & & & & \\ \vdots & \ddots & \ddots & \vdots & & & & & \\ 0 & & \ddots & 1 & & & & & \\ 0 & \cdots & \cdots & \lambda_1 & & & & & \\ & & & & \lambda_2 & 1 & \cdots & 0 & \\ & & & & \vdots & \ddots & \ddots & \vdots & \\ & & & & 0 & & \ddots & 1 & \\ & & & & 0 & \cdots & \cdots & \lambda_2 & \\ & & & & & & & & \ddots \end{pmatrix}.$$

Quando denotamos a k-ésima coluna de T por t_k $(k = 1, 2, \ldots, n)$, substituímos a equação matricial

$$AT = TJ$$

pelo sistema equivalente de equações

$$At_1 = \lambda_1 t_1, At_2 = \lambda_1 t_2 + t_1, \ldots, At_{p_1} = \lambda_1 t_{p_1} + t_{p_1-1}$$

$$At_{p_1+1} = \lambda_2 t_{p_1+1}, At_{p_1+2} = \lambda_2 t_{p_1+2} + t_{p_1+1}, \ldots, At_{p_1+p_2} = \lambda_2 t_{p_1+p_2} + t_{p_1+p_2-1} \ldots$$

as quais podemos escrever como:

$$(A - \lambda_1 I)t_1 = 0, (A - \lambda_1 I)t_2 = t_1, \ldots, (A - \lambda_1 I)t_{p_1} = t_{p_1-1}$$

$$(A - \lambda_2 I)t_{p_1+1} = 0, (A - \lambda_2 I)t_{p_1+2} = t_{p_1+1}, \ldots, (A - \lambda_2 I)t_{p_1+p_2} = t_{p_1+p_2-1} \ldots$$

Todas as colunas de T são divididas em *Cadeias de Jordan* de colunas:

$$[t_1, t_2, \ldots, t_{p_1}], [t_{p_1+1}, t_{p_1+2}, \ldots, t_{p_1+p_2}], \ldots$$

Para cada Bloco de Jordan de J (ou, o que significa a mesma coisa, cada divisor elementar $(\lambda - \lambda_1)^{p_1}, \ldots$) há uma Cadeia de Jordan de colunas. Cada Cadeia de Jordan de colunas é caracterizada por um sistema de equações do tipo das definidas acima.

O trabalho de encontrar uma matriz transformação T se reduz a encontrar as Cadeias de Jordan que darão as n colunas linearmente independentes.

Mostraremos que estas Cadeias de Jordan de colunas podem ser determinadas através da matriz adjunta reduzida $C(\lambda)$.

Para a matriz $C(\lambda)$ nós temos a identidade

$$(\lambda I - A)C(\lambda) = \psi(\lambda)I,$$

onde $\psi(\lambda)$ é o polinômio minimal de A.

Seja

$$\psi(\lambda) = (\lambda - \lambda_0)^m \chi(\lambda); \quad (\chi(\lambda_0) \neq 0).$$

Derivamos a identidade $(\lambda I - A)C(\lambda) = \psi(\lambda)I$, termo a termo, $m - 1$ vezes:

$$(\lambda I - A)C'(\lambda) + C(\lambda) = \psi'(\lambda)I$$
$$(\lambda I - A)C''(\lambda) + 2C'(\lambda) = \psi''(\lambda)I$$
$$(\lambda I - A)C'''(\lambda) + 3C'(\lambda) = \psi'''(\lambda)I$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$(\lambda I - A)C^{(m-1)}(\lambda) + (m-1)C^{(m-2)}(\lambda) = \psi'^{(m-1)}(\lambda)I$$

Substituindo λ_0 por λ temos:

$$(\lambda I - A)C(\lambda) = \psi(\lambda)I.$$

$$(\lambda I - A)C(\lambda) = (\lambda - \lambda_0)^m \chi(\lambda)I$$

$$(\lambda I - A)C(\lambda) = 0 \Rightarrow (a - \lambda_0 I)C = 0.$$

$$(\lambda_0 I - A)C'(\lambda_0) + C(\lambda_0 = \psi'(\lambda_0)I = 0$$

$$(\lambda_0 I - A)C'(\lambda_0) = -C(\lambda_0) \Rightarrow (A - \lambda_0 I)D = C,$$

onde $D = \frac{C'(\lambda_0)}{1!}$.

$$(\lambda_0 I - A)C''(\lambda_0) + 2C'(\lambda_0) = 0$$

$$(\lambda_0 I - A)C''(\lambda_0) = -2C'(\lambda_0) \Rightarrow (A - \lambda_0 I)C''(\lambda_0) = 2D \Rightarrow (A - \lambda_0 I)E = D,$$

onde $E = \frac{1}{2!}C''(\lambda_0)$.

Fazendo o mesmo para todas as equações acima, temos:

$$(A - \lambda_0 I)K = G,$$

onde $G = \frac{1}{(m-2)!}C^{m-2}(\lambda_0)$ e $K = \frac{1}{(m-1)!}C^{m-1}(\lambda_0)$.

Daí, temos para as k-ésimas colunas $(k = 1, 2, \ldots, n)$:

$$(A - \lambda_0 I)C_{1j} = 0, (A - \lambda_0 I)C_{2j} = C_{1k}, \ldots, (A - \lambda_0 I)C_{kj} = C_{k-1,j}(j = 1, \ldots, n).$$

Sendo $C = C(\lambda_0) \neq 0$ (pois caso $C(\lambda_0) = 0$, todos os elementos de $C(\lambda)$, teriam um divisor comum de grau positivo), podemos encontrar um $j(\leq n)$ tal que

$$C_{1j} \neq 0.$$

Então as m colunas

$$C_{1j}, C_{2j}, C_{3j}, \ldots, C_{k-1,j}, C_{kj}$$

são linearmente independentes, ou seja, existem $\gamma, \delta, \ldots, \chi \in \mathbb{F}$, tais que

$$\gamma C_{1,j} + \delta C_{2,j} + \ldots + \chi C_{k,j} = 0$$

se e somente se

$$\gamma = \delta = \ldots = \chi = 0.$$

Caso as colunas $C_{1j}, C_{2j}, \ldots, C_{kj}$ sejam linearmente independentes satisfaçam o sistema de equações

$$(A - \lambda_0 I)C_{1j} = 0, (A - \lambda_0 I)C_{2j} = C_{1j}, \ldots, (A - \lambda_0 I)C_{kj} = C_{k-1,j}, \quad (j = 1, \ldots, n)$$

elas formam uma Cadeia de Jordan de vetores correspondentes aos divisores elementares $(\lambda - \lambda_0)^m$.

Se $C_{1j} = 0$ para algum j, mas $C_{2j} \neq 0$, então as colunas $C_{2j}, \ldots, C_{k-1,j}, C_{kj}$ formam uma Cadeia de Jordan de $m - 1$ vetores e assim por diante.

Veremos agora como construir a matriz transformação T no caso de os divisores elementares de A serem primos entre si.

$$(\lambda - \lambda_1)^{(m_1)}, \ldots, (\lambda - \lambda_s)^{(m_s)}, (\lambda_i \neq \lambda_j \ \ para \ \ i \neq j; i, j = 1, 2, \ldots, s).$$

Aos divisores elementares $(\lambda - \lambda_j)^{m_j}$ nós associamos a Cadeia de Jordan de colunas

$$C^{(j)}, D^{(j)}, \ldots, G^{(j)}, K^{(j)},$$

construída como mostrado anteriormente. Então

$$(A - \lambda_j I)C^{(j)} = 0, (A - \lambda_j I)D^{(j)} = C^{(j)}, \ldots, (A - \lambda_j I)K^{(j)} = G^{(j)}.$$

Quando atribuímos valores a j $(1, 2, \ldots, s)$, obtemos s Cadeias de Jordan contendo n colunas ao todo. Essas colunas são linearmente independentes.

Supomos que

$$\sum_{j=1}^{s} [\gamma_j C^{(j)} + \delta_j D^{(j)} + \ldots + \rho_j K^{(j)}] = 0.$$

Multiplicamos ambos os lados, pela esquerda, por $(A - \lambda_1 i)^{m_1} \dots (A - \lambda_{j-1} i)^{m_{j-1}}$ $(A - \lambda_j i)^{m_j} (A - \lambda_{j+1} i)^{m_{j+1}} \dots (A - \lambda_s i)^{m_s}$ e obtemos

$$\rho_j = 0.$$

Substituindo $m_j - 1$, sucessivamente, por $m_j - 2$, $m_j - 3, \dots$, temos

$$\gamma_j = \delta_j = \dots = \rho_j = 0 \quad (j = 1, 2, \dots, s).$$

Definimos então a matriz T pela fórmula

$$T = (C^{(1)}, D^{(1)}, \dots, K^{(1)}; C^{(2)}, D^{(2)}, \dots, K^{(2)}, \dots, C^{(s)}, D^{(s)}, \dots, K^{(s)}).$$

Capítulo 5

Operador linear em um espaço n-dimensional (teoria geométrica de divisores elementares)

5.1 O polinômio minimal de um vetor e de um espaço

Consideramos um espaço vetorial n-dimensional V sobre $\mathbb{C}$ e um operador linear A neste espaço.

Seja x um vetor arbitrário de V. Formamos a seqüência de vetores

$$x, Ax, A^2x, \ldots .$$

Sendo V um espaço de dimensão finita, existe $p \in \mathbb{Z}(1 \leq p \leq n)$ tal que os vetores $x, Ax, \ldots, A^{p-1}x$ são linearmente independentes e A^px é uma combinação linear destes vetores com coeficientes em $\mathbb{C}$:

$$A^px = -\gamma_1 A^{p-1}x - \gamma_2 A^{p-2}x - \ldots - \gamma_p x.$$

Tomamos o polinômio mônico [1] $\varepsilon(\lambda) = \lambda^p + \gamma_1\lambda^{p-1} + \ldots + \gamma_{p-1}\lambda + \gamma_p$. Então temos: $\varepsilon(A)x = A^px + \gamma_1 A^{p-1}x + \ldots + \gamma_{p-1}Ax + \gamma_p x = -\gamma_1 A^{p-1}x - \ldots - \gamma_{p-1}Ax - \gamma_p x + \gamma_1 A^{p-1}x + \ldots + \gamma_{p-1}Ax + \gamma_p x = 0$.

Todo polinômio $\varepsilon(\lambda)$ no qual $\varepsilon(A)x = 0$ é chamado de polinômio anulador do vetor x em relação ao operador A dado. Construímos o polinômio mônico anulador de x com menor grau possível. Então podemos chamá-lo de polinômio anulador minimal

[1] polinômio mônico é um polinômio no qual o coeficiente do termo de maior grau é igual a 1

de x ou simplesmente polinômio minimal do vetor x.

Todo polinômio anulador $\varepsilon(\lambda)$ de x é divisível pelo polinômio minimal $\psi(\lambda)$, ou seja,

$$\varepsilon(\lambda) = \psi(\lambda)\alpha(\lambda) + \varrho(\lambda),$$

onde $\alpha(\lambda)$, $\varrho(\lambda)$ são o quociente e o resto, respectivamente, na divisão de $\varepsilon(\lambda)$ por $\psi(\lambda)$. Então

$$0 = \varepsilon(A)x = \alpha(A)\psi(A)x + \varrho(A)x = \alpha(A).0 + \varrho(A)x = \varrho(A)x.$$

Isso implica que $\varrho(A)x = 0$. Mas o grau de $\varrho(\lambda)$ é menor que o grau de $\psi(\lambda)$. Então $\varrho(\lambda) \equiv 0$. Disso segue que: *Todo vetor x tem um único polinômio minimal.*

Escolhemos uma base, $e_1, e_2, \ldots, e_n$, de V. Então todo vetor x de V pode ser escrito como

$$X = x_1e_1 + x_2e_2 + \ldots + x_ne_n.$$

Denotamos por $\psi_1(\lambda)$, $\psi_2(\lambda)$,..., $\psi_n(\lambda)$ os polinômios minimais associados aos vetores da base ($e_1, e_2, \ldots, e_n$, respectivamente) e por $\varepsilon(\lambda)$ o mínimo múltiplo comum destes polinômios. Temos que o polinômio $\varepsilon(\lambda)$ é mônico e é também um polinômio anulador dos vetores da base.

Segue que:

$$\varepsilon(A)x = \varepsilon(A)[x_1e_1 + x_2e_2 + \ldots + x_ne_n] =$$

$$x_1\varepsilon(A)e_1 + x_2\varepsilon(A)e_2 + \ldots + x_n\varepsilon(A)e_n = 0 + 0 + \ldots + 0 = 0,$$

ou seja, $\varepsilon(A)$ se anula em todo vetor $x \in V$. Então

$$\varepsilon(A) = 0.$$

O polinômio $\varepsilon(\lambda)$ é chamado de *Polinômio anulador para todo o espaço V.*

Seja $\psi(\lambda)$ um polinômio anulador arbitrário para todo o espaço V. Então $\varepsilon(\lambda)$ é um polinômio anulador para os vetores da base, $e_1, e_2, \ldots, e_n$. Então $\psi(\lambda)$ é um múltiplo comum dos polinômios minimais $\psi_1(\lambda), \psi_2(\lambda), \ldots, \psi_n(\lambda)$ destes vetores e, portanto, é divisível sem resto pelo mínimo múltiplo comum $\varepsilon(\lambda)$. Então $\varepsilon(\lambda)$ tem grau mínimo e é mônico. Esse polinômio é determinado unicamente pelo espaço V e pelo operador A, e é chamado de polinômio minimal do espaço V. A unicidade do polinômio minimal do espaço V vem do fato que: *todo polinômio anulador $\varepsilon(\lambda)$ do espaço V é divisível pelo polinômio minimal $\psi(\lambda)$.* Apesar da construção do polinômio minimal ter sido associada a uma base finita, este polinômio não depende da escolha desta base.

Pelo fato de que o polinômio minimal de V anula todo o vetor x de V, temos que o polinômio minimal do espaço é divisível pelo polinômio minimal de todo vetor do espaço.

Chamamos a atenção para o fato de que todas as afirmações feitas neste capítulo são baseadas em um operador A o qual foi estabelecido anteriormente. Alterando o operador, podemos também estar alterando os polinômios.

5.2 Decomposição em subespaços invariantes com polinômios minimais primos entre si

Definição 30 *Sejam V' e V'' dois subespaços vetoriais do espaço V. Dizemos que V é decomposto em dois subespaços, V' e V'', se são verdadeiras as seguintes condições:*

1. *V' e V'' não possuem vetores em comum, exceto o vetor nulo;*
2. *todo vetor $x \in V$ pode ser representado pela forma*

$$x = x' + x''$$

onde $x' \in V'$ e $x'' \in V''$.

Podemos representar tal decomposição por

$$V = V' \oplus V''.$$

Pelo condição 1, temos que a representação de x como soma de subespaços é feita de maneira única. Podemos verificar isto da seguinte forma: Representamos x de duas formas distintas.

$$x = x' + x''; x' \in V' \text{ e } x'' \in V''$$

$$x = y' + y''; y' \in V' \text{ e } y'' \in V''.$$

Subtraindo uma equação da outra temos

$$x - x = x' + x'' - y' - y'',$$

ou seja,

$$x' - y' = y'' - x''.$$

Mas $x' - y' \in V'$ e $y'' - x'' \in V''$. Pela condição 1, $x' - y' = 0$ e $y'' - x'' = 0$, pois o único vetor comum de V e V' é o vetor nulo. Logo, $x' = y'$ e $x'' = y''$.

Definição 31 *Um subespaço $V' \subset V$ é chamado de invariante em relação ao operador A se $AV' \subset V'$, ou seja, se $x \in V'$ então $Ax \in V'$.*

A definição acima nos diz que o operador A leva um vetor de um subespaço invariante em um vetor do mesmo subespaço.

Teorema 20 *(Primeiro teorema na decomposição de um espaço em subespaços invariantes) Se para um dado operador A, o polinômio minimal $\psi(\lambda)$ do espaço é representado, sobre $\mathbb{C}$, na forma de produto de dois polinômios, $\psi_1(\lambda)$ e $\psi_2(\lambda)$, primos entre si e com coeficiente do termo de maior grau igual a 1, ou seja, $\psi(\lambda) = \psi_1(\lambda)\psi_2(\lambda)$, então todo o espaço V é dividido em dois subespaços invariantes, I_1 e I_2, isto é,*

$$V = I_1 \oplus I_2,$$

onde $\psi_1(\lambda)$ e $\psi_2(\lambda)$ são os polinômios minimais de I_1 e I_2, respectivamente.

Demonstração:

Definimos por I_1 o conjunto de todos os vetores $x \in V$ que satisfazem a equação $\psi_1(A)x = 0$ e, por I_2, o conjunto de todos os vetores $x \in V$ que satisfazem a equação $\psi_2(A)x = 0$. I_1 e I_2 são subespaços de V.

Assumimos que ψ_1 e ψ_2 são primos entre si. Então existem polinômios com coeficientes em $\mathbb{C}$, $\chi_1(\lambda)$ e $\chi_2(\lambda)$, tais que:

$$\psi_1(\lambda)\chi_1(\lambda) + \psi_2(\lambda)\chi_2(\lambda) = 1.$$

Tomamos um vetor arbitrário x de V. Substituímos λ por A na equação acima e aplicamos os dois lados da mesma em x:

$$x = \psi_2(A)\chi_2(A)x + \psi_1(A)\chi_1(A)x,$$

ou seja,

$$x = x' + x'',$$

onde $x' = \psi_2(A)\chi_2(A)x$ e $x'' = \psi_1(A)\chi_1(A)x$. Além disso,

$$\psi_1(A)x' = \psi_1\psi_2(A)\chi_2(A)x = \psi(A)\chi_2(A)x = 0$$

e

$$\psi_2(A)x'' = \psi_2\psi_1(A)\chi_1(A)x = \psi(A)\chi_1(A)x = 0.$$

Isso implica que

$$x' \in I_1 \quad \text{e} \quad x'' \in I_2.$$

Se tomarmos $x_0 \in I_1 \cap I_2$, ou seja, $\psi_1(A)x_0 = 0$ e $\psi_2(A)x_0 = 0$ então

$$x_0 = \chi_1(A)\psi_1(A)x_0 + \chi_2(A)\psi_2(A)x_0 = 0 + 0 = 0.$$

Logo I_1 e I_2 têm apenas o vetor nulo em comum. Portanto,

$$V = I_1 \oplus I_2.$$

Supomos que $x \in I_1$. Então $\psi_1(A)x = 0$, ou ainda, $A\psi_1(A)x = A.0 = 0$. Isso significa que $\psi_1(A)Ax = 0$. Portanto, $Ax \in I_1$. Disso podemos concluir que I_1 é um subespaço invariante. Da mesma forma provamos que I_2 também é um subespaço invariante.

Para completarmos a demonstração do teorema, precisamos mostrar que $\psi_1(\lambda)$ e $\psi_2(\lambda)$ são polinômios minimais de I_1 e I_2, respectivamente.

Sejam $\psi_1'(\lambda)$ um polinômio anulador de I_1 e x um vetor unitário de V. Mas, podemos escrever x como

$$x = x' + x''$$

onde $x' \in I_1$ e $x'' \in I_2$. Então

$$\psi_1'(A)\psi_2(A)x = \psi_1'(A)\psi_2(A)x' + \psi_1'(A)\psi_2(A)x'' =$$

$$\psi_2(A)\psi_1'(A)x' + \psi_1'(A)\psi_2(A)x''0 + 0 = 0.$$

Sendo x um vetor arbitrário de V, então o produto $\psi_1'(\lambda)\psi_2(\lambda)$ é um polinômio anulador de V e, portanto, é divisível por $\psi(\lambda)$ sem resto. Mas, $\psi(\lambda) = \psi_1(\lambda)\psi_2(\lambda)$. Logo, $\psi_1'(\lambda)$ é divisível por $\psi_1(\lambda)$. Como $\psi_1(\lambda)$ divide todos os polinômios anuladores, e o polinômio minimal ψ_1 de I_1 é um polinômio anulador, temos que $\psi_1(\lambda)$ divide $\psi_1'(\lambda)$. Mas $\psi_1'(\lambda)$ divide $\psi_1(\lambda)$. Como os dois são mônicos, podemos afirmar que $\psi_1'(\lambda) = \psi_1(\lambda)$.

Entretanto, $\psi_1'(\lambda)$ é um polinômio anulador arbitrário de I_1 e $\psi_1(\lambda)$ é um polinômio particular dos polinômios anuladores, pela definição de I_1. Portanto, $\psi_1(\lambda)$ é um polinômio minimal de I_1. Do mesmo modo provamos que $\psi_2(\lambda)$ é polinômio minimal de I_2.

Isso completa a demonstração do teorema.

Decompomos $\psi(\lambda)$ em fatores irredutíveis sobre $\mathbb{C}$:

$$\psi(\lambda) = [\lambda - \psi_1]^{c_1}[\lambda - \psi_2]^{c_2} \dots [\lambda - \psi_s]^{c_s},$$

onde $[\lambda - \psi_i]^{c_i}$, $(i = 1, 2, \dots, s)$, são polinômios irredutíveis distintos sobre $\mathbb{C}$. Além disso, esses polinômios são mônicos. Então, pelo teorema 20, temos

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_s,$$

onde I_k é um subespaço invariante com polinômio minimal $[\lambda - \psi_k]^{c_k}$ $(k = 1, 2, \ldots, s)$.

Lema 1 *Se os polinômios minimais dos vetores e' e e'' são primos entre si, então o polinômio minimal do vetor soma $e' + e''$ é igual ao produto dos polinômios minimais dos vetores e' e e''.*

Demonstração:

Sejam $\chi_1(\lambda)$ e $\chi_2(\lambda)$ os polinômios minimais dos vetores e' e e'', respectivamente. Pela hipótese, $\chi_1(\lambda)$ e $\chi_2(\lambda)$ são primos enter si. Seja $\chi(\lambda)$ um polinômio anulador arbitrário do vetor e, onde $e = e' + e''$. Então

$$0 = \chi_2(A)\chi(A)e = \chi_2(A)\chi(A)e' + \chi_2(A)\chi(A)e''$$

$$\chi_2(A)\chi(A)e' = \chi_2(A)\chi(A)e - \chi_2(A)\chi(A)e'' = 0 - 0 = 0$$

Logo, $\chi_2(\lambda)\chi(\lambda)$ é um polinômio anulador de e'. Portanto, $\chi_2(\lambda)\chi(\lambda)$ é divisível por $\chi_1(\lambda)$. Mas, como $\chi_1(\lambda)$ e $\chi_2(\lambda)$ são primos entre si, podemos concluir que $\chi(\lambda)$ é divisível por $\chi_2(\lambda)$. Logo, $\chi(\lambda)$ é divisível por $\chi_1(\lambda)\chi_2(\lambda)$, ou seja, todo polinômio anulador de e é divisível por $\chi_1(\lambda)\chi_2(\lambda)$. Portanto, $\chi_1(\lambda)\chi_2(\lambda)$ é o polinômio minimal do vetor e.

Isso completa a demonstração.

Teorema 21 *Em um espaço vetorial sempre existe um vetor no qual seu polinômio minimal coincide com o polinômio minimal de todo o espaço.*

Demonstração:

Consideramos o caso em que o polinômio minimal do espaço V é uma potência do polinômio linear $\varphi(\lambda)$, ou seja

$$\psi(\lambda) = [\lambda - \varphi]^l.$$

Tomamos $e_1, e_2, \ldots, e_n$ como base de V. O polinômio minimal de e_1 é um divisor de $\psi(\lambda)$ e é representado pela forma $[\lambda - \varphi]^{l_i}$, onde $l_i \leq l(i = 1, 2, \ldots, n)$. Mas o polinômio minimal do espaço é o mínimo múltiplo comum dos polinômios minimais dos vetores da base e $\psi(\lambda)$ é a maior das potências $[\lambda - \varphi]^{l_i}$, $(i = 1, 2, \ldots, n)$. Isso significa que $\psi(\lambda)$ é igual ao polinômio minimal de um dos vetores da base.

Passaremos agora para o caso geral. Escrevemos o espaço como soma de subespaços invariantes:

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_s,$$

cujos polinômios minimais são $(\lambda - \varphi_1)^{l_1}$, $(\lambda - \varphi_2)^{l_2}$, $\ldots$, $(\lambda - \varphi_s)^{l_s}$, respectivamente. Existem vetores $e_1 \in I_1$, $e_2 \in I_2, \ldots, e_s \in I_s$ cujos polinômios minimais são os polinômios minimais de I_j. Pelo lema 1, o polinômio minimal do vetor $e = e_1 + e_2 + \ldots + e_s$ é igual ao produto

$$[\lambda - \varphi_1]^{l_1}[\lambda - \varphi_2]^{l_2} \ldots [\lambda - \varphi_s]^{l_s},$$

ou seja, é igual ao polinômio minimal do espaço V.

5.3 Congruência e Espaços quocientes

Definição 32 *Seja I um subespaço vetorial tal que $I \subset V$. Dizemos que dois vetores $x, y \in V$ são congruentes módulo I (notação: $x \equiv y(mod\ I)$) se e somente se $y - x \in I$.*

Vejamos algumas propriedades:

P 20 $x \equiv x(mod\ I)$. *(reflexão)*

Demonstração:

$$x \equiv x(mod\ I) \Leftrightarrow x - x \in I \Leftrightarrow 0 \in I$$

Como I é um subespaço vetorial, podemos afirmar que $0 \in I$.

P 21 $x \equiv y(mod\ I) \Rightarrow y \equiv x(mod\ I)$ *(simetria)*

Demonstração:

$$x \equiv y(mod\ I) \Rightarrow y - x \in I$$

Como I é um subespaço vetorial, temos que $(-1)(y - x) \in I$.

$$x - y \in I \Rightarrow y \equiv x(mod\ I).$$

P 22 $x \equiv y(mod\ I)$ *e* $y \equiv z(mod\ I) \Rightarrow x \equiv z(mod\ I)$ *(transitividade)*

Demonstração:

$$x \equiv y(mod\ I) \Rightarrow y - x \in I$$

$$y \equiv z(mod\ I) \Rightarrow z - y \in I$$

Como I é um subespaço vetorial, temos que

$$y - x + z - y \in I.$$

$$z - x \in I.$$

Logo, $x \equiv z(mod\ I)$.

A partir destas três propriedades passaremos a utilizar a congruência para dividir todos os vetores do espaço em classes. Os pares de vetores congruentes $(mod\ I)$ pertencem à mesma classe. A classe que contém o vetor x será denotada por $\overline{x}$. O subespaço I é uma dessas classes, chamada de $\overline{o}$.

As congruências podem ser somadas termo a termo e multiplicadas por um escalar pertencente a $\mathbb{C}$:

P 23 $x \equiv x'(mod\ I)$ e $y \equiv y'(mod\ I) \Rightarrow x + y \equiv x' + y'(mod\ I)$

Demonstração:

$$x \equiv x'(mod\ I) \Rightarrow x' - x \in I$$

$$y \equiv y'(mod\ I) \Rightarrow y' - y \in I$$

Sendo I um subespaço vetorial, temos

$$x' - x + y' - y \in I \Rightarrow x' + y' - (x + y) \in I \Rightarrow x + y \equiv x' + y'(mod\ I).$$

P 24 $x \equiv x'(mod\ I) \Rightarrow \alpha x \equiv \alpha x'(mod\ I), \alpha \in \mathbb{C}$.

Demonstração:

$$x \equiv x'(mod\ I) \Rightarrow x' - x \in I$$

Sendo I um subespaço vetorial, temos

$$\alpha(x' - x) \in I,$$

onde $\alpha \in \mathbb{F}$. Logo,

$$\alpha x' - \alpha x \in I \Rightarrow \alpha x \equiv \alpha x'(mod\ \ I).$$

Por essas propriedades, temos que as operações de adição e multiplicação por um escalar não alteram as classes de congruência. A soma dos vetores desta classe é denotado por $\overline{x} + \overline{y}$ e a multiplicação por escalar por $\alpha\overline{x}$.

Denotamos por $\overline{V}$ o conjunto de todas as classes $\overline{x}$, $\overline{y}, \ldots$, nas quais a adição e a multiplicação por escalar são definidas. Então, tanto V quanto $\overline{V}$ são espaços vetoriais sobre $\mathbb{C}$. Dizemos que $\overline{V}$ é um espaço quociente de V. Se n, m, $\overline{n}$ são as dimensões dos espaços V, I, $\overline{V}$, respectivamente, então $\overline{n} = n - m$.

Sejam A um operador linear em V e I um subespaço invariante em relação à A. Isso significa que, $AI \subset I$, ou seja, se $x \in I$ então $Ax \in AI$. Para $x \equiv x'(mod\ \ I)$ temos que $x' - x \in I$. Logo, $A(x' - x) \in AI$, ou ainda, $Ax' - Ax \in AI$. Daí vem que $Ax \equiv Ax'(mod\ \ AI)$. Concluímos então que o operador A pode ser aplicado aos dois lados da congruência. Em outras palavras, se o operador A é aplicado a todos os vetores $x, x', \ldots$ da classe $\overline{x}$, então os vetores $Ax, Ax', \ldots$ também pertencem a uma classe, a qual denotamos por $A\overline{x}$. O operador linear A leva uma classe em outra classe e preserva as operações, portanto, é um operador linear em $\overline{V}$.

Definição 33 *Dizemos que os vetores $x_1, x_2, \ldots, x_p$ são linearmente dependentes módulo I se existem números $\alpha_1, \alpha_2, \ldots, \alpha_p$ pertencentes a $\mathbb{C}$, não simultaneamente nulos, tais que*

$$\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p \equiv 0(mod\ \ I).$$

Não só o conceito de dependência linear, mas todos os conceitos, afirmações e conseqüências, podem ser repetidas, palavra por palavra, apenas substituindo $' ='$ por $' \equiv (mod\ \ I)'$, onde I é um subespaço invariante fixado em relação a A.

5.4 Decomposição de um espaço em subespaços cíclicos invariantes

Sejam V um espaço vetorial, A um operador linear e um vetor de V. Seja $\sigma(\lambda) = \lambda^p + \alpha_1\lambda^{p-1} + \ldots + \alpha_{p-1}\lambda + \alpha_p$ o polinômio minimal de e. Então os vetores

$$e, Ae, \ldots, A^{p-1}e$$

são linearmente independentes e

$$A^p e = -\alpha_p e - \alpha_{p-1}Ae - \ldots - \alpha_1 A^{p-1}e.$$

Os vetores $e, Ae, \ldots, A^{p-1}e$ formam uma base p-dimensional do subespaço I. Chamamos este subespaço de cíclico considerando a condição especial da base formada por esses vetores e do vetor A^pe.

Tomamos novamente os vetores $e, Ae, \ldots, A^{p-1}e$. O operador A leva o primeiro destes vetores no segundo, o segundo no terceiro, e assim por diante. O último vetor da base é levado, também por A, para a combinação linear dos vetores da base, A^pe. Assim, A leva todos os vetores da base em vetores de I e um vetor arbitrário de I em outro vetor arbitrário de I. Em outras palavras, um subespaço cíclico é sempre invariante em relação à A.

Todo vetor $x \in I$ é representado como combinação linear dos vetores da base, ou seja,

$$x = \chi(A)e,$$

onde $\chi(\lambda)$ é um polinômio em λ de grau menor ou igual a $p-1$ com coeficientes em $\mathbb{C}$. Construindo todos os polinômios possíveis, $\chi(\lambda)$, de grau menor ou igual $p-1$ com coeficientes em $\mathbb{C}$, obtemos todos os vetores de I, mais ainda, polinômios diferentes produzem vetores diferentes. Considerando a base acima citada e a fórmula $x = \chi(A)e$, dizemos que o vetor e gera o subespaço.

Teorema 22 *(Segundo teorema na decomposição de um espaço em subespaços invariantes) Em relação a um dado operador linear A dado, o espaço vetorial V sempre pode ser decomposto numa soma direta de subespaços cíclicos $I_1, I_2, \ldots, I_t$ com polinômios minimais $\psi_1(\lambda), \psi_2(\lambda), \ldots, \psi_t(\lambda)$, tais que $\psi_1(\lambda)$ coincide com o polinômio minimal $\psi(\lambda)$ do espaço todo e cada $\psi_i(\lambda)$ é divisível por $\psi_{i+1}(\lambda)$, $(i = 1, 2, \ldots, t-1)$.*

Demonstração:

Sejam $\psi_1(\lambda) = \psi(\lambda) = \alpha^m + \alpha_1\lambda^{m-1} + \ldots + \alpha_m$ o polinômio minimal do espaço vetorial V. Então, pelo teorema 21, existe um vetor e no espaço para o qual este polinômio é minimal. Denotamos por I_1 o subespaço cíclico com base

$$e, Ae, \ldots, A^{m-1}e.$$

Se $m = n$ então $V = I_1$. Supomos que $n > m$ e que o polinômio

$$\psi_2(\lambda) = \lambda^p + \beta_1\lambda^{p-1} + \ldots + \beta_p$$

é o polinômio minimal de $V(mod\ I)$. Pelo o que vimos anteriormente, $\psi_2(\lambda)$ é um divisor de $\psi_1(\lambda)$, ou seja, existe um polinômio $x(\lambda)$ tal que

$$\psi_1(\lambda) = \psi_2(\lambda)x(\lambda) = x(\lambda)\psi_2(\lambda).$$

Entretanto, em V existe um vetor g^* no qual o polinômio minimal relativo é $\psi_2(\lambda)$. Então

$$\psi_2(A)g^* \equiv 0,$$

ou seja, existe um polinômio $\chi(\lambda)$ de grau menor ou igual a $m-1$ tal que

$$\psi_2(A)g^* \equiv \chi(A)e.$$

Aplicamos o operador $x(A)$ em ambos os lados da equação. Então temos

$$\psi_2(A)g^* = \chi(A)e$$

$$x(A)\psi_2(A)g^* = x(A)\chi(A)e$$

$$\psi_1(A)g^* = x(A)\chi(A)e.$$

Mas $\psi_1(\lambda)$ é o polinômio minimal absoluto do espaço. Então

$$x(A)\chi(A)e = 0.$$

Logo $x(\lambda)\chi(\lambda)$ é um polinômio anulador do vetor e e, portanto, divisível pelo polinômio minimal $\psi_1(\lambda) = x(\lambda)\psi_2(\lambda)$. Daí, $\chi(\lambda)$ é divisível por $\psi_2(\lambda)$, ou seja,

$$\chi(\lambda) = x_1(\lambda)\psi_2(\lambda) = \psi_2(\lambda)x_1(\lambda),$$

onde $x_1(\lambda)$ é um polinômio. Então temos

$$\psi_2(A)g^* = \chi(A)e$$

$$\psi_2(A)g^* = \psi_2(A)x_1(A)e$$

$$\psi_2(A)g^* - \psi_2(A)x_1(A)e = 0$$

$$\psi_2(A)[g^* - x_1(A)e] = 0.$$

Definimos $g = g^* - x_1(A)$. Portanto

$$\psi_2(A)g = 0.$$

Então $\psi_2(\lambda)$ é um polinômio anulador absoluto do vetor g e, portanto, é divisível pelo polinômio minimal absoluto de g. Por outro lado,

$$g = g^* - x_1(A) \Rightarrow g \equiv g^*(mod\ \ I_1).$$

Sendo $\psi_2(\lambda)$ o polinômio minimal relativo de g^*, este também é para g. Então $\psi_2(\lambda)$ é, simultaneamente, polinômio minimal absoluto e relativo de g. Pelo fato de $\psi_2(\lambda)$ ser o polinômio minimal absoluto de g, temos que o subespaço I_2 com base

$$g, Ag, \ldots, A^{p-1}g$$

é cíclico.

Sendo $\psi_2(\lambda)$ o polinômio minimal relativo de $g(mod\ I)$, temos que os vetores $g, Ag, \ldots, A^{p-1}g$ são linearmente independentes $(mod\ I_1)$. Podemos então afirmar a independência linear de $m+p$ vetores

$$e, Ae, \ldots, A^{m-1}e; g, Ag, \ldots, A^{p-1}g.$$

Estes vetores formam uma base do subespaço invariante $I_1 + I_2$ de dimensão $m+p$. Se $n = m+p$ então $V = I_1 + I_2$. Se $n > m+p$, consideramos $V(mod\ I_1 + I_2)$ e continuamos nosso processo de separação em subespaços cíclicos. Sendo V um espaço de dimensão finita (n), este processo se encerrará para algum subespaço I_t, onde $t \leq n$. Isso prova o teorema.

Teorema 23 *Um espaço é cíclico se, e somente se, sua dimensão é igual ao grau de seu polinômio minimal.*

Demonstração:

Sejam V um espaço cíclico n-dimensional e $\psi(\lambda) = \lambda^m + \alpha_1\lambda^{m-1} + \ldots + \alpha_{m-1}\lambda + \alpha_m$ seu polinômio minimal. Então, pela definição de espaço cíclico, $m = n$.

Sejam V um espaço vetorial arbitrário n-dimensional e m o grau de seu polinômio minimal. Pela hipótese, $m = n$. Pelo teorema da decomposição, V pode ser representado na forma

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_t.$$

Mas a dimensão do subespaço cíclico I_1 é m, pois seu polinômio minimal coincide com o polinômio minimal do espaço todo. Como $m = n$, temos $V = I_1$, ou seja, V é um espaço cíclico.

Teorema 24 *Um espaço cíclico pode ser dividido em subespaços invariantes que*

1. *também são cíclicos;*

2. *possuem polinômios minimais primos entre si.*

Demonstração:

Decompomos o espaço cíclico V em dois subespaços invariantes I_1 e I_2:

$$V = I_1 \oplus I_2.$$

Denotamos as dimensões de V, I_1 e I_2 por n, n_1 e n_2, seus polinômios minimais por $\psi(\lambda)$, $\psi_1(\lambda)$ e $\psi_2(\lambda)$, e os graus destes polinômios por m, m_1 e m_2, respectivamente. Então $m_1 \leq n_1$ e $m_2 < n_2$. Somando estas inequações, termo a termo:

$$m_1 + m_2 \leq n_1 + n_2.$$

Sendo $\psi(\lambda)$ o mínimo múltiplo comum entre $\psi_1(\lambda)$ e $\psi_2(\lambda)$, temos:

$$m \leq m_1 + m_2.$$

Entretanto, $n = n_1 + n_2$. Logo,

$$m \leq m_1 + m_2 \leq n_1 + n_2 = n.$$

Mas, como V é cíclico, temos que $m = n$. Então

$$m = m_1 + m_2 = n_1 + n_2 = n.$$

O fato de termos $m = m_1 + m_2$, nos diz que $\psi_1(\lambda)$ e $\psi_2(\lambda)$ são primos entre si. Temos ainda que: $m_1 \leq n_1$, $m_2 \leq n_2$ e $m_1 + m_2 = n_1 + n_2$. Então $m_1 = n_1$ e $m_2 = n_2$. Portanto, os subespaços I_1 e I_2 são cíclicos.

Teorema 25 *Se um espaço é dividido em subespaços invariantes que:*

1. *são cíclicos;*

2. *possuem polinômios minimais primos entre si*

então o espaço é cíclico.

Demonstração:

Dividimos o espaço vetorial V em dois subespaços invariantes cíclicos, I_1 e I_2,

$$V = I_1 \oplus I_2.$$

Sejam $\psi(\lambda)$, $\psi_1(\lambda)$ e $\psi_2(\lambda)$ os polinômios minimais de V, I_1 e I_2, m, m_1 e m_2 os graus destes polinômios e n, n_1 e n_2 as dimensões de V, I_1 e I_2, respectivamente. Pela hipótese, $\psi_1(\lambda)$ e $\psi_2(\lambda)$ são primos entre si. Então, pelo lema 1

$$\psi(\lambda) = \psi_1(\lambda)\psi_2(\lambda).$$

Logo,

$$m = m_1 + m_2.$$

Sendo I_1 e I_2 subespaços cíclicos, pelo teorema 23, temos que $m_1 = n_1$ e $m_2 = n_2$. Mas $n = n_1 + n_2$. Logo,

$$n = n_1 + n_2 = m_1 + m_2 = m.$$

Sendo $n = m$, temos que o espaço é cíclico. Isso completa a demonstração.

Teorema 26 *Um espaço não pode ser dividido em subespaços próprios invariantes se e somente se*

1. *é cíclico*

2. *seu polinômio minimal é potência de um polinômio linear.*

Demonstração:

Seja V um espaço que não pode ser dividido em subespaços invariantes. Então, V é um espaço cíclico pois, caso contrário, pelo segundo teorema da decomposição, V poderia ser dividido em subespaços cíclicos.

O polinômio minimal de V é uma potência de um polinômio irredutível pois, caso não o fosse, pelo primeiro teorema da decomposição, V poderia ser dividido em subespaços invariantes.

Sejam V um espaço cíclico e seu polinômio minimal uma potência de um polinômio linear, ou seja,

$$\psi(\lambda) = [\lambda - \varphi]^c.$$

Neste caso, o polinômio minimal de todos os subespaços invariantes de V também é uma potência de seu polinômio irredutível $\varphi(\lambda)$. Então, os polinômios minimais de dois subespaços invariantes quaisquer, não são primos entre si. Logo, V não pode ser dividido em subespaços invariantes. Isso completa a demonstração do teorema.

Teorema 27 *(Terceiro teorema na decomposição de um espaço em subespaços invariantes): Um espaço sempre pode ser dividido em subespaços cíclicos invariantes*

$$V = I' \oplus I'' \oplus \ldots \oplus I^{(u)}$$

tal que o polinômio minimal de cada subespaço cíclico é uma potência de um polinômio irredutível.

Demonstração:

Dividimos o espaço vetorial V em subespaços cíclicos:

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_t.$$

Dividimos também os polinômios minimais destes subespaços em fatores lineares:

$$\psi_1(\lambda) = [\lambda - \varphi_1]^{c_{11}}[\lambda - \varphi_2]^{c_{12}} \ldots [\lambda - \varphi_s]^{c_{1s}},$$

$$\psi_2(\lambda) = [\lambda - \varphi_1]^{c_{21}}[\lambda - \varphi_2]^{c_{22}} \ldots [\lambda - \varphi_s]^{c_{2s}},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\psi_t(\lambda) = [\lambda - \varphi_1]^{c_{t1}}[\lambda - \varphi_2]^{c_{t2}} \ldots [\lambda - \varphi_s]^{c_{ts}},$$

$$(c_{ik} \geq c_{kj};\ \ i, k = 1, 2, \ldots, t;\ \ j = 1, 2, \ldots, s).$$

Aplicamos o primeiro teorema da decomposição para I_1:

$$I_1 = I_1' + I_1'' + I_1^{(s)};$$

onde I_1', I_1'', $I_1^{(s)}$ são subespaços cíclicos cujos polinômios minimais são $[\lambda - \varphi_1]^{c_{11}}, [\lambda - \varphi_2]^{c_{12}}, \ldots, [\lambda - \varphi_s]^{c_{1s}}$. Da mesma forma, decompomos I_2, ..., I_t. Assim obtemos a decomposição de V em subespaços cíclicos com polinômios minimais $[\lambda - \varphi_1]^{c_{1j}}, [\lambda - \varphi_2]^{c_{2j}}, \ldots, [\lambda - \varphi_k]^{c_{tj}}$, $(j = 1, 2, \ldots, s)$.

Assim, fica provado o teorema.

5.5 A forma canônica de um operador

Seja I_1 um subespaço invariante m-dimensional do espaço V. Tomamos uma base para I_1, e_1, e_2,..., e_m, e completamos esta para formarmos uma base para V:

$$e_1, e_2, \ldots, e_m, e_{m+1}, \ldots, e_n.$$

Denotamos por B a matriz do operador A nesta base. Vimos anteriormente que a k-ésima coluna de B é formada pelas coordenadas do vetor Ae_k $(k = 1, 2, \ldots, n)$. Para $k \leq m$, o vetor $Ae_k \in I_1$ e as últimas $n - m$ coordenadas de Ae_k são zero. Por tanto, B tem a seguinte forma:

$$\begin{pmatrix} B_1 & B_3 \\ 0 & B_2 \end{pmatrix},$$

onde B_1 e B_2 são matrizes quadradas de ordem m e $n-m$, respectivamente, e B_3 é uma matriz retangular de ordem $m \times (n-m)$. O fato do quarto *bloco* ser zero, mostra a invariância do subespaço I_1 (com relação a base e_1, e_2 ,..., e_m).

Assumimos que $e_{m+1}, \ldots, e_n$ é a base de algum subespaço invariante I_2 tal que $V = I_1 \oplus I_2$ e as bases dos subespaços invariantes, I_1 e I_2, formam uma base para S. Então o bloco B_3 também é zero e a matriz B tem a forma quase-diagonal:

$$\begin{pmatrix} B_1 & 0 \\ 0 & B_2 \end{pmatrix} = \{B_1, B_2\},$$

onde B_1 e B_2 são, respectivamente, matrizes quadradas de ordem m e $n-m$, para as quais foram dados operadores nos subespaços I_1 e I_2 (em relação as bases $e_1, e_2, \ldots, e_m$ e $e_{m+1}, \ldots, e_n$). Da mesma forma, uma matriz quase-diagonal corresponde à decomposição do espaço em subespaços invariantes.

Pelo segundo teorema da decomposição, podemos dividir o espaço V em subespaços cíclicos I_1, I_2,..., I_t:

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_t.$$

Cada um dos polinômios minimais destes subespaços, $\psi_1(\lambda)$, $\psi_2(\lambda)$,..., $\psi_1(\lambda)$, é divisor de seu antecessor.

Sejam

$$\psi_1 = \lambda^m + \alpha_{11}\lambda^{m-1} + \ldots + \alpha_{1m},$$

$$\psi_2 = \lambda^p + \alpha_{21}\lambda^{p-1} + \ldots + \alpha_{2p},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\psi_t = \lambda^v + \alpha_{t1}\lambda^{v-1} + \ldots + \alpha_{tv},$$

onde $(m \geq p \geq \ldots \geq v)$.

Denotamos por e_1, e_2, ..., e_t, os vetores que geram os subespaços I_1, I_2, ..., I_t e formamos uma base para o espaço a partir das bases dos subespaços cíclicos:

$$e_1, Ae_2, \ldots, A^{m-1}e_1; e_2, Ae_2, \ldots, A^{p-1}e_2, \ldots; l, Ae_t, \ldots, A^{v-1}e_t.$$

A matriz L_I corresponde ao operador A na base dada. Então a matriz L_I tem a forma quase-diagonal:

$$\begin{pmatrix} L_1 & 0 & \cdots & o \\ 0 & L_2 & \cdot & o \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & L_t \end{pmatrix}.$$

A matriz L_1 corresponde ao operador A em I_1 em relação à base $b_1 = e_1$, $b_2 = Ae_1, \ldots, b_m = A^{m-1}e_1$.

$$L_1 = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_{1m} \\ 1 & 0 & \cdots & 0 & -\alpha_{1,m-1} \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & -\alpha_{12} \\ 0 & 0 & \cdots & 1 & -\alpha_{11} \end{pmatrix}.$$

Da mesma forma:

$$L_2 = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_{2p} \\ 1 & 0 & \cdots & 0 & -\alpha_{2,p-1} \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & -\alpha_{22} \\ 0 & 0 & \cdots & 1 & -\alpha_{21} \end{pmatrix}.$$

Calculando os polinômios característicos das matrizes $L_1, L_2, \ldots, L_t$, encontramos:

$$\det(\lambda I - L_1) = \psi_1(\lambda), \det(\lambda I - L_2) = \psi_2(\lambda), \ldots, \det(\lambda I - L_y) = \psi_t(\lambda).$$

Para subespaços cíclicos o polinômio característico de um operador A coincide com o polinômio minimal do subespaço relativo a este operador.

Então a matriz L_I corresponde ao operador A na base canônica. Se B é a matriz correspondente a A em uma base arbitrária, então B é semelhante a L_I, ou seja, existe uma matriz não singular T tal que

$$B = TL_IT^{-1}.$$

Dizemos que a matriz L_I tem a *Primeira Forma Canônica Natural.* Essa forma é caracterizada por:

1. sua forma quase-diagonal;

2. sua estrutura especial de blocos diagonais;

3. o polinômio característico de cada bloco é divisível pelo polinômio característico do bloco seguinte.

Do mesmo modo, pelo terceiro teorema da decomposição, obtemos uma matriz L_{II} relacionada ao operador A em uma base apropriada. Essa matriz tem a *Segunda Forma Canônica Natural*, que é caracterizada por:

1. sua forma quase-diagonal

$$L_{II} = \{L^{(1)}, L^{(2)}, \ldots, L^{(u)}\};$$

2. sua estrutura especial de blocos diagonais;

3. o polinômio característico de cada bloco diagonal é uma potência de um polinômio linear.

5.6 Polinômios Invariantes e Divisores Elementares

Denotamos por $D_p(\lambda)$ o maior divisor comum de todos os menores de ordem p da matriz característica $B_\lambda = \lambda I - B$, $(p = 1, 2, \ldots, n)$. Sendo cada polinômio da seqüência

$$D_n(\lambda), D_{n-1}(\lambda), \ldots, D_1(\lambda)$$

divisível pelo seu sucessor, as fórmulas

$$i_1(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}, i_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \ldots, i_n(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)}; (D_o(\lambda) \equiv 1)$$

definem n polinômios cujos produtos é igual ao polinômio característico

$$\Delta(\lambda) = \det(\lambda I - B) = D_n(\lambda) = i_1(\lambda) i_2(\lambda) \ldots i_n(\lambda).$$

Dividimos os polinômios $i_p(\lambda)$, $(p = 1, 2, \ldots, n)$ em fatores lineares:

$$i_p(\lambda) = (\lambda - \varphi_1)^{\gamma_p} (\lambda - \varphi_2)^{\delta_p} \ldots . \quad (p = 1, 2, \ldots, n);$$

onde $\varphi_1(\lambda)$, $\varphi_2(\lambda)$,... são polinômios lineares distintos. Os polinômios $i_1(\lambda), i_2(\lambda), \ldots, i_n(\lambda)$ são chamados de polinômios invariantes, e todas as potências não-constantes $(\lambda - \varphi_1)^{\gamma_p}$, $(\lambda - \varphi_2)^{\delta_2}$,... são chamadas de divisores elementares da matriz característica $B_\lambda = \lambda I - B$ ou, simplesmente, de B.

Tanto o produto de todos os divisores elementares, como o produto de todos os polinômios invariantes, é igual ao polinômio característico $\Delta(\lambda) = \det(\lambda I - B)$.

O nome *polinômio invariante* é justificado pelo fato de que duas matrizes semelhantes B e $\tilde{B}$, tais que

$$\tilde{B} = T^{-1}BT,$$

sempre têm os mesmos polinômios invariantes

$$i_p(\lambda) = \tilde{i}_p(\lambda) \quad (p = 1, 2, \ldots, n).$$

Temos então que

$$\tilde{B}_\lambda = \lambda I - \tilde{B} = T^{-1}(\lambda I - B)T = T^{-1}B_\lambda T.$$

Pela fórmula de Binet-Cauchy, há uma relação entre os menores das matrizes semelhantes B_λ e $\tilde{B}_\lambda$:

$$\tilde{B}_\lambda \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} =$$

$$\sum_{\substack{\alpha_{11} < \alpha_{12} < \cdots < \alpha_{1p} \\ \alpha_{21} < \alpha_{22} < \cdots < \alpha_{2p}}} T^{-1} \begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1p} \end{pmatrix} A_\lambda \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1p} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2p} \end{pmatrix} T \begin{pmatrix} \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2p} \\ k_1 & k_2 & \ldots & k_p \end{pmatrix},$$

onde $(p = 1, 2, \ldots, n)$.

Esta equação mostra que todo divisor comum dos menores de ordem p e B_λ é também, um divisor comum dos menores de ordem p de $\tilde{B}_\lambda$, e vice-versa. Deve-se a isso o fato de que

$$D_p(\lambda) = \tilde{B}_p(\lambda) \;\; e \;\; i_p(\lambda) = \tilde{i}_p(\lambda) \;\; (p = 1, 2, \ldots, n).$$

Portanto, todas as matrizes que representam um operador A em bases distintas, são semelhantes e por isso possuem os mesmos polinômios invariantes e os mesmos divisores elementares.

Teorema 28 *(Forma mais precisa do segundo teorema da decomposição) Se A é um operador linear em V, então o espaço pode ser decomposto em subespaços cíclicos*

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_p$$

tais que na seqüência de polinômios minimais $\psi_1(\lambda), \psi_2(\lambda), \ldots, \psi_p(\lambda)$ dos subespaços I_1, $I_2, \ldots$, I_p, cada um é divisível pelo seguinte. Os polinômios minimais são determinados de forma única: eles coincidem com os polinômios invariantes, diferentes de

1, do operador A.

Demonstração:

Seja V um espaço vetorial e A um operador linear em V. Pelo teorema 27, podemos dividir V em subespaços cíclicos invariantes, I_1, I_2, $\ldots$, I_p, ou seja:

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_p,$$

onde $\psi_i(\lambda)$ é o polinômio minimal do subespaço I_i, $(i = 1, 2, \ldots, p)$. Escolhemos bases para os subespaços de maneira que a matriz L_i, $(i = 1, 2, \ldots, p)$, que representa o operador A, tenha a forma L_I, ou seja, a primeira forma canônica natural. Seja B uma matriz de blocos diagonais:

$$B = \begin{pmatrix} L_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & L_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & L_s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & L_t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & L_p \end{pmatrix}.$$

Inicialmente iremos considerar que retiramos uma coluna de L_s e uma linha de L_t. Assim obtemos matrizes retangulares $\hat{L}_s$ e $\hat{L}_t$.

Seja C a matriz $\begin{pmatrix} L_s & 0 & 0 & 0 & 0 \\ 0 & L_{s+1} & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & L_{t-1} & 0 \\ 0 & 0 & 0 & 0 & L_t \end{pmatrix}$. Tomamos por $\hat{C}$ a matriz C com remoção de uma linha de L_t e uma coluna de L_s. Formamos então uma nova matriz $\hat{B}$ com a seguinte forma:

$$\hat{B} = \begin{pmatrix} L_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{C} & 0 & 0 & 0 \\ 0 & 0 & 0 & L_t & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & L_p \end{pmatrix},$$

onde $\det \hat{B} = \det L_1 \ldots . \det L_{s-1} . \det \hat{C} . \det L_{t+1} \ldots \det L_p$. Afirmamos que $\hat{C}$ contém uma linha ou coluna de zeros e, portanto, é zero. Lembramos que C é uma matriz diagonal com $L_s \dots L_t$ na diagonal.

Dividimos $\hat{C}$ em quatro blocos,

$$\begin{pmatrix} U & 0 \\ T & V \end{pmatrix},$$

sendo U a matriz composta pelas primeiras n_{s-1} linhas e colunas e V o bloco complementar de U na diagonal. Como $\hat{C}$ é *diagonal* o bloco superior à direita de U é formado por zeros. Deste modo, $\det \hat{C} = \det U. \det V$. Olhamos agora para a primeira linha de V. Observamos que L_s é de ordem n_s e L_{s+1} começa na posição C_{n_s+1,n_s+1}, mas em $\hat{C}$ uma coluna foi removida em L_s. Assim U é de ordem $n_s - 1$ e termina com o elemento C_{n_s-1,n_s}.

Deste modo V começa com $C_{n_s,n_{s+1}}$. Como C é diagonal por blocos e $C_{n_s,n_{s+1}}$ está à direita do bloco L_s, $C_{n_s,n_{s+j}} = 0$, para todo $j > 0$ e a primeira linha de V é nula.

Se temos uma linha de L_s e uma coluna de L_t, a demonstração é semelhante, mas V terá uma coluna de zeros.

Como $\det V = 0$, $\det \hat{C} = \det U. \det V = 0$ e $\det \hat{M} = 0$.

Supomos agora que tiramos uma linha e uma coluna do mesmo bloco diagonal L_s. Assim,

$$\hat{B} = \begin{pmatrix} L_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{L}_s & 0 & 0 & 0 \\ 0 & 0 & 0 & L_{s+1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & L_p \end{pmatrix},$$

Assim $\det \hat{B} = \det L_1 \dots \det L_{s-1} \det L_{s+1} \det \hat{L}_s$. Portanto

$$\frac{\det \hat{B}}{\det L_2 \dots \det L_p} = \frac{\det L_1 \dots \det L_{s-1} \det L_{s+1} \dots \det L_p \det \hat{L}_s}{\det L_2 \dots \det L_p} = \frac{\det L_1}{\det L_s} \det \hat{L}_s$$

e como $\det L_s$ divide $\det L_1$, temos que $\det L_2 \dots \det L_p$ divide $\det B$. Desta forma $\det L_2 \dots \det L_p$, divide todos os cofatores de B e é um múltiplo do MDC dos cofatores, pois essa é uma característica da matriz L_I.

Agora construímos o cofator de M onde eliminamos a linha 1 e a coluna n_s, a última coluna de L_1. Deste modo L_1 se torna triangular com -1 em toda a diagonal. Daí $\det \hat{B} = (-1)^{n_s-1}. \det L_2 \dots \det L_p$, ou seja, $\det L_2 \dots \det L_p$ é o MDC dos cofatores.

Isso completa a demonstração.

Esse teorema pode ser escrito das seguintes maneiras:

* Para cada operador linear A em V existe uma base na qual a matriz L_I deste

mesmo operador é da primeira forma canônica natural. Esta matriz é unicamente determinada quando o operador A é dado: os polinômios característicos dos blocos diagonais de L_I são os polinômios invariantes de V.

* Em cada classe de matrizes semelhantes existe uma matriz L_I que tem a primeira forma canônica natural. Os polinômios característicos dos blocos diagonais L_I coincidem com os polinômios invariantes, além de 1, das matrizes desta classe.

Teorema 29 *Duas matrizes escalares são semelhantes se, e somente se, elas têm os mesmos polinômios invariantes.*

Demonstração:

Já vimos anteriormente que duas matrizes semelhantes possuem os mesmos polinômios invariantes.

Sejam B e C duas matrizes escalares que possuem os mesmos polinômios invariantes. Sendo a matriz L_I determinada de maneira única quando estes polinômios são dados, as matrizes B e C são semelhantes à matriz L_I. Portanto, B e C são semelhantes.

Teorema 30 *Se A é um operador linear em um espaço vetorial V, então V pode ser dividido em subespaços cíclicos onde os polinômios minimais são divisores elementares de A.*

Demonstração:

O polinômio característico $\Delta(\lambda)$ do operador A coincide com $D_n(\lambda)$, e portanto com o produto de todos os polinômios invariantes:

$$\Delta(\lambda) = \psi_1(\lambda)\psi_2(\lambda)\ldots\psi_t(\lambda).$$

Mas $\psi_1(\lambda)$ é o polinômio minimal do espaço com relação à A; logo $\psi_1(A) = 0$ e, pela equação acima,

$$\Delta(A) = 0.$$

Disso obtemos o teorema de Hamilton-Cayley. Pela divisão dos polinômios $\psi_1(\lambda)$, $\psi_2(\lambda),\ldots,\ \psi_t(\lambda)$ em fatores lineares:

$$\psi_1(\lambda) = [\lambda - \varphi_1]^{c_{11}}[\lambda - \varphi_2]^{c_{12}} \dots [\lambda - \varphi_s]^{c_{1s}}$$

$$\psi_2(\lambda) = [\lambda - \varphi_1]^{c_{21}}[\lambda - \varphi_2]^{c_2} \dots [\lambda - \varphi_s]^{c_{2s}}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\psi_t(\lambda) = [\lambda - \varphi_1]^{c_{t1}}[\lambda - \varphi_2]^{c_{t2}} \dots [\lambda - \varphi_s]^{c_{ts}}$$

$$(c_{ik} \geq c_{kj}, \quad i, k = 1, 2, \dots, t; \quad j = 1, 2, \dots, s),$$

chegamos ao terceiro teorema da decomposição. Para cada potência com expoente diferente de zero do lado direito das equações acima há um subespaço invariante correspondente a esta decomposição. Então todas as potências, diferentes de 1, entre $[\lambda - \varphi_k]^{c_{1k}}, \dots, [\lambda - \varphi_k]^{c_{tk}}$, $(k = 1, 2, \dots, s)$ são divisores elementares de A no conjunto C.

Isso prova o teorema.

Seja $V = I_1 \oplus I_2 \oplus \dots \oplus I_u$ uma decomposição do espaço V. Denotamos por e_1, e_2,..., e_u os vetores que geram os subespaços I_1, I_2,..., I_u e a partir das bases destes subespaços formamos a base do espaço

$$e_1, Ae_2, \dots, e_2, Ae_2, \dots; e_u, Ae_u, \dots.$$

A matriz L_{II} corresponde aos operador A em relação à base acima tem a forma quase-diagonal, como L_I:

$$L_{II} = \{L_1, L_2, \dots, L_u\}.$$

Os blocos diagonais $L_1, L_2, \dots, L_u$ possuem a mesma estrutura de blocos apresentados na página 72 (L_1 e L_2). Entretanto, os polinômios característicos destes blocos diagonais não são polinômios invariantes, mas sim os divisores elementares de A. A matriz L_{II} tem a segunda forma canônica natural.

Isso nos dá uma outra formulação para o teorema 30:

Teorema 30′ *Para cada operador linear A em V existe uma base na qual a matriz L_{II} do operador dado tem a segunda forma canônica natural; os polinômios característicos dos blocos diagonais são divisores elementares de A.*

Esse teorema também admite uma formulação em termos de matrizes:

Teorema 30″ *Uma matriz A com elementos em $\mathbb{C}$ é sempre semelhante a uma matriz L_{II} a qual tem a segunda forma canônica natural na qual os polinômios carac-*

terísticos dos blocos diagonais são divisores elementares de A.

Teorema 31 *Se o espaço vetorial V é dividido em subespaços invariantes (em relação a um operador A) que podem ser decompostos, então os polinômios minimais destes subespaços são divisores elementares de A.*

Demonstração:

Seja $V = I_1 \oplus I_2 \oplus \ldots \oplus I_u$ uma decomposição arbitrária de um espaço V em subespaços invariantes que não podem ser decompostos. Então, pelo teorema 26, os subespaços I_1, $I_2 \ldots I_u$ são cíclicos e seus polinômios minimais são potências de polinômios lineares. Podemos escrever estas potências na forma:

$$[\lambda - \varphi_1]^{c_{11}}, [\lambda - \varphi_2]^{c_{12}}, \ldots, [\lambda - \varphi_s]^{c_{1s}},$$

$$[\lambda - \varphi_1]^{c_{21}}, [\lambda - \varphi_2]^{c_{22}}, \ldots, [\lambda - \varphi_s]^{c_{2s}},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$[\lambda - \varphi_1]^{c_{t1}}, [\lambda - \varphi_2]^{c_{t2}}, \ldots, [\lambda - \varphi_s]^{c_{ts}},$$

$$(c_{ik} \geq c_{kj}, \quad i, k = 1, 2, \ldots, t; \ j = 1, 2, \ldots, s).$$

Denotamos por I_1 a soma dos subespaços cujos polinômios minimais formam a primeira das equações acima. Fizemos o mesmo com I_2, ..., I_t, onde t é o número de linhas das equações acima. Pelo teorema 25, os subespaços I_1, I_2, ..., I_t são cíclicos e seus polinômios minimais $\psi_1(\lambda)$, $\psi_2(\lambda)$, ..., $\psi_t(\lambda)$ são determinados pelas fórmulas:

$$\psi_1(\lambda) = [\lambda - \varphi_1]^{c_{11}}.[\lambda - \varphi_2]^{c_{12}}.\ldots.[\lambda - \varphi_s]^{c_{1s}},$$

$$\psi_2(\lambda) = [\lambda - \varphi_1]^{c_{21}}.[\lambda - \varphi_2]^{c_{22}}.\ldots.[\lambda - \varphi_s]^{c_{2s}},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\psi_t(\lambda) = [\lambda - \varphi_1]^{c_{t1}}.[\lambda - \varphi_2]^{c_{t2}}.\ldots.[\lambda - \varphi_s]^{c_{ts}},$$

$$(c_{ik} \geq c_{kj}, \quad i, k = 1, 2, \ldots, t; \ j = 1, 2, \ldots, s).$$

Na seqüência $\psi_1(\lambda)$, $\psi_2(\lambda)$, ..., $\psi_t(\lambda)$ cada polinômio é divisível pelo seguinte. Podemos então aplicar o teorema 28 para a decomposição

$$V = I_1 \oplus I_2 \oplus \ldots \oplus I_t.$$

Logo,

$$\psi_p(\lambda) = i_p(\lambda) \quad (p = 1, 2, \ldots, n),$$

e portanto, todas as potências $[\lambda - \varphi_k]^{c_k}$, $(k = 1, 2, \ldots, s)$ com expoentes diferentes de zero são os divisores elementares de A no conjunto $\mathbb{C}$.

Isso prova o teorema.

Há uma formulação equivalente em termos de matrizes:

Teorema 31' *Em cada classe de matrizes semelhantes existe uma única matriz (dentro dos blocos diagonais) que possue a segunda forma canônica L_{II}; os polinômios característicos destes blocos diagonais são divisores elementares de todas as matrizes da classe dada.*

Teorema 32 *Se o espaço V é dividido em subespaços invariantes com relação ao operador A, então os divisores elementares de A de cada subespaço invariante formam um sistema completo de divisores elementares de A em V.*

Demonstração:

Supomos que o espaço V é dividido em dois subespaços invariantes (em relação ao operador A)

$$V = I_1 \oplus I_2.$$

Quando dividimos I_1 e I_2 em subespaços que não podem ser decompostos, obtemos a decomposição de V em subespaços que não podem ser decompostos. Pelo teorema 31 fica completa a demonstração deste teorema.

Este teorema tem a seguinte formulação para matrizes a qual é usada para encontrar os divisores elementares de uma matriz.

Teorema 32' *Um sistema completo de divisores elementares de uma matriz quase-diagonal é obtido pela união dos divisores elementares dos blocos diagonais.*

5.7 A forma canônica de Jordan de um operador

Supomos que todos as raízes do polinômio característico $\Delta(\lambda)$ de um operador A pertencem à $\mathbb{C}$.

Neste caso, a decomposição dos polinômios invariantes em divisores elementares em $\mathbb{C}$ será vista da seguinte maneira:

$$i_1(\lambda) = [\lambda - \varphi_1]^{c_{11}}.[\lambda - \varphi_2]^{c_{12}}.\ldots.[\lambda - \varphi_s]^{c_{1s}},$$

$$i_2(\lambda) = [\lambda - \varphi_1]^{c_{21}}.[\lambda - \varphi_2]^{c_{22}}.\ldots.[\lambda - \varphi_s]^{c_{2s}},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$i_t(\lambda) = [\lambda - \varphi_1]^{c_{t1}}.[\lambda - \varphi_2]^{c_{t2}}.\ldots.[\lambda - \varphi_s]^{c_{ts}},$$

$$(c_{ik} \geq c_{kj}, \quad i, k = 1, 2, \ldots, t; \;\; j = 1, 2, \ldots, s).$$

Sendo o produto de todos os polinômios invariantes igual ao polinômio característico $\Delta(\lambda)$, podemos afirmar que λ_1, λ_2, ..., λ_s são raízes distintas de $\Delta(\lambda)$.

Tomamos um divisor elementar arbitrário

$$(\lambda - \lambda_0)^p;$$

onde λ_0 é uma das raízes do polinômio característico e p é um dos expoentes (diferente de zero) c_k, d_k,..., l_k, $(k = 1, 2, \ldots, s)$.

Para este divisor elementar há um subespaço cíclico I correspondente, que é gerado por um vetor, o qual iremos denotar por e. Para este vetor $(\lambda - \lambda_0)^p$ é o polinômio minimal.

Consideramos os vetores

$$e_1 = (A - \lambda_0 I)^{p-1}e, e_2 = (A - \lambda_0 I)^{p-2}e \ldots, e_p = e.$$

Os vetores e_1, e_2, $\ldots e_p$ são linearmente independentes. Por outro lado, existe um polinômio anulador para e de grau menor que p e isso é impossível de acontecer. Notamos que

$$(A - \lambda_0 I)e_1 = 0, (A - \lambda_0 I)e_2 = e_1, \ldots, (A - \lambda_0 I)e_p = e_{p-1}$$

$$Ae_1 = \lambda_0 e_1, Ae_2 = \lambda_0 e_2 + e_1, \ldots, Ae_p = \lambda_0 e_p + e_{p-1}.$$

A partir disso podemos escrever a matriz correspondente a A em I para a base e_1, e_2, ..., e_p:

$$\lambda_0 I^{(p)} + H^{(p)} = \begin{pmatrix} \lambda_0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & & \ddots & 1 \\ 0 & \cdots & \cdots & \lambda_0 \end{pmatrix},$$

onde $I^{(p)}$ é a matriz identidade de ordem p e $H^{(p)}$ é a matriz de ordem p que possui $1's$ na sua primeira diagonal superior e $0's$ nas demais posições.

Os vetores independentes e_1, e_2, ..., e_p para os quais as últimas equações apresentadas são verdadeiras formam a Cadeia de Jordan de vetores em I. A Cadeia de Jordan relacionada com cada subespaço I_1, I_2, ..., I_u, forma uma base de Jordan de V. Se denotarmos os polinômios minimais destes subespaços, ou seja, os divisores elementares de A, por

$$(\lambda-\lambda_1)^{p_1}, (\lambda-\lambda_2)^{p_2}, \ldots, (\lambda-\lambda_u)^{p_u}\ ^2,$$

então a matriz J correspondente a A na base de Jordan tem a seguinte forma quase-diagonal:

$$J = \{\lambda_1 I^{(p_1} + H^{(p_1)}, \lambda_2 I^{(p_2} + H^{(p_2)}, \ldots, \lambda_u I^{(p_u)} + H^{(p_u)}\}.$$

Dizemos que a matriz J possue a *Forma Canônica de Jordan*, ou simplesmente, *Forma de Jordan*. A matriz J pode ser escrita imediatamente quando os divisores elementares de A no conjunto $\mathbb{C}$ que contêm todas as raízes características da equação $\Delta(\lambda) = 0$ são conhecidos.

Toda matriz B é semelhante a uma matriz J com Forma Canônica de Jordan, ou seja, para uma matriz arbitrária B sempre existe uma matriz T não-singular tal que

$$B = TJT^{-1}.$$

Se todos os divisores elementares de A são de grau 1 então a matriz de Jordan é uma matriz diagonal e temos:

$$B = T\{\lambda_1, \lambda_2, \ldots, \lambda_n\}T^{-1}.$$

Definimos agora os vetores e_1, e_2, ..., e_p, os quais já foram definidos anteriormente, de maneira inversa:

$$g_1 = e_p = e, g_2 = e_{p-1} = (A - \lambda_0 I)e, \ldots, g_p = e_1 = (A - \lambda_0 I)^{p-1}e.$$

Então

$$(A - \lambda_0 I)g_1 = g_2, (A - \lambda_0 I)g_2 = g_3, \ldots, (A - \lambda_0 I)g_p = 0.$$

Segue que

$$Ag_1 = \lambda_0 g_1 + g_2, Ag_2 = \lambda_0 g_2 + g_3, \ldots, Ag_p = \lambda_0 g_p.$$

Os vetores g_1, g_2, ..., g_p formam uma base para o subespaço cíclico invariante I que correspondem aos divisores elementares $(\lambda - \lambda_0)^p$.

[2] os números λ_1, λ_2, ..., λ_u não precisam ser todos distintos

Nesta base, temos uma matriz correspondente ao operador A:

$$\lambda_0 I^{(p)} + F^{(p)} = \begin{pmatrix} \lambda_0 & 0 & 0 & \cdots & 0 \\ 1 & \lambda_0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 1 & \lambda_0 \end{pmatrix}$$

Dizemos que os vetores g_1, g_2, $\ldots$, g_p formam a Cadeia Inferior de Jordan dos vetores. Se tomarmos essa cadeia de vetores em cada subespaço I', I'', $\ldots$, $I^{(u)}$, formamos uma base inferior de Jordan na qual o operador A corresponde a matriz quase-diagonal

$$J_1 = \{\lambda_1 I^{(p_1)} + F^{(p_1)}, \lambda_2 I^{(p_2)} + F^{(p_2)}, \ldots, \lambda_u I^{(p_u)} + F^{(p_u)}\}.$$

Dizemos que a matriz J_1 é da Forma Inferior de Jordan. Chamaremos a forma que vimos anteriormente de Forma Superior de Jordan.

Podemos concluir que:

Toda matriz é semelhante a uma matriz inferior e a uma matriz superior de Jordan.

Capítulo 6

Uma aplicação da forma canônica da Jordan

Consideramos um sistema de equações lineares homogêneas de primeira ordem com coeficientes constantes:

$$\begin{cases} \frac{dx_1}{dt} = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \frac{dx_2}{dt} = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{dx_n}{dt} = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n \end{cases}$$

onde t é a variável independente, x_1, x_2, $\ldots$, x_n são funções desconhecidas de t e a_{ik} $(i = 1, 2, \ldots, n)$ são números complexos.

Seja A a matriz quadrada, $A = (a_{ik})_1^n$, dos coeficientes das equações e $x = (x_1, x_2, \ldots, x_n)$ uma matriz coluna. Podemos escrever o sistema acima na forma de uma única equação diferencial matricial

$$\frac{dx}{dt} = Ax.$$

Chamamos de derivada de uma matriz a matriz obtida a partir da substituição dos elementos da matriz dada por suas derivadas. Então $\frac{dx}{dt}$ é a matriz coluna cujos elementos $\frac{dx_1}{dt}$, $\frac{dx_2}{dt}$, $\ldots$, $\frac{dx_n}{dt}$.

Procuramos uma solução do sistema de equações diferenciais que satisfaça as seguintes condições iniciais:

$$x_1\big|_{t=0} = x_{10},\ x_2\big|_{t=0} = x_{20},\ x_n\big|_{t=0} = x_{n0},$$

ou ainda,

$$x\big|_{t=0} = x_0.$$

Expandimos a coluna x pela série de MacLaurin em potências de t:

$$x = x_0 + x_0't + x_0''\frac{t^2}{2!} + \ldots (x_0' = \frac{dx}{dt}|_{t=0},\ x_0'' = \frac{d^2x}{dt^2}|_{t=0}, \ldots).$$

Derivando sucessivamente a equação $\frac{dx}{dt} = Ax$, temos:

$$\frac{d^2x}{dt^2} = A\frac{dx}{dt} = A^2x,\quad \frac{d^3x}{dt^3} = A^3x, \ldots.$$

Substituindo t por 0, temos:

$$x_0' = Ax_0,\quad x_0'' = A^2x_0, \ldots.$$

Podemos então escrever a série acima da seguinte maneira:

$$x = x_0 + tAx_0 + \frac{t^2}{2!}A^2x_0 + \ldots + e^{AT}x_0.$$

Temos que $\frac{d}{dt}(e^{At}) = \frac{d}{dt}(I + At + \frac{A^2t^2}{2!} + \ldots) = A + A^2t + \frac{A^3t^2}{2!} + \ldots = Ae^{At}$ e $\frac{dx}{dt} = Ax$. Então $x = e^{At}x_0$ é solução da equação diferencial $\frac{dx}{dt} = Ax$. Para $t = 0$, temos

$$x = e^{At}x_0 = e^0x_0 = x_0,$$

ou seja,

$$x|_{t=0} = x_0.$$

Há grande interesse em funções do tipo e^A, $sen(A)$ e $cos(A)$.

Seja f uma função analítica com série de potência

$$f(x) = \alpha_0 + \alpha_1x + \ldots + \alpha_nx^n + \ldots.$$

Definimos

$$f(A) = \alpha_0I + \alpha_1A + \ldots + \alpha_nA^n + \ldots.$$

Para $e^A = I + A + \frac{1}{2!}A^2 + \ldots + \frac{1}{n!}A^n + \ldots$ temos como tarefa calcular o limite desta seqüência. Sabemos que este trabalho não pode ser considerado fácil.

Entretanto, $A = TJT^{-1}$ onde T é uma matriz inversível e J uma matriz de Jordan. Assim

$$f(A) = \alpha_0I + \alpha_1A + \ldots + \alpha_nA^n + \ldots =$$

$$\alpha_0(TIT^{-1}) + \alpha_1(TJT^{-1}) + \ldots + \alpha_n(TJT^{-1})^n + \ldots =$$

$$\alpha_0 TIT^{-1} + \alpha_1 TJT^{-1} + \ldots + \alpha_n TJ^nT^{-1} + \ldots =$$

$$T(\alpha_0 I + \alpha_1 J + \ldots + \alpha_n J^n + \ldots)T^{-1}.$$

Dessa forma, precisamos apenas calcular $f(J)$. Mas $J = J_1 \oplus J_2 \oplus \ldots \oplus J_s$ e todo J_i é da forma $\lambda_i I + H$ onde λ_i é um autovalor de A e H é uma matriz que possui 1 na sua diagonal e zero nas demais posições.

Desse modo $J^n = J_1^n \oplus J_2^n \oplus \ldots \oplus J_s^n = (\lambda_i I + H_1)^n + \ldots + (\lambda_s I + H_s)^n$, ou seja, basta analisar as potências de Jordan. Supomos que $dim\,(J_i) = n = dim\, H$. Pela forma de H, verificamos que H^2 tem 1 na primeira sobrediagonal e zeros nas outras posições. Em geral, para $1 \leq j \leq n-1$, H^j tem 1 na j-ésima sobrediagonal e zeros nas demais posições; para $j \geq n$, $H^j = 0$. Deste modo,

$$(J_i)^k = (\lambda I)^k + \binom{k}{1}(\lambda I)^{k-1}H + \ldots.$$

Temos que:

$$J^0 = I = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

$$J^1 = \lambda I + H = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 \\ 0 & \lambda & 1 & \ldots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \ldots & \ldots & \lambda \end{pmatrix}$$

$$J^2 = (\lambda I + H)^2 = \lambda^2 + 2\lambda + H^2 = \begin{pmatrix} \lambda^2 & 2\lambda & 1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda^2 & 2\lambda & 1 & 0 & \ldots & 0 \\ 0 & 0 & \lambda^2 & 2\lambda & 1 & \ddots & 0 \\ \vdots & \vdots & & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & & & \ddots & \ddots & 1 \\ \vdots & \vdots & & & & \ddots & 2\lambda \\ 0 & 0 & \ldots & \ldots & \ldots & \ldots & \lambda^2 \end{pmatrix}$$

$J^3 = (\lambda I + H)^3 = \lambda^3 + 3\lambda^2 H + 3\lambda H^2 + H^3 =$

$$\begin{pmatrix} \lambda^3 & 3\lambda^2 & 3\lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda^3 & 3\lambda^2 & 3\lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda^3 & 3\lambda^2 & 3\lambda & 1 & \ddots & 0 \\ \vdots & \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & & & \ddots & \ddots & \ddots & 1 \\ \vdots & \vdots & & & & \ddots & \ddots & 3\lambda \\ \vdots & \vdots & & & & & \ddots & 3\lambda^2 \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & \lambda^3 \end{pmatrix}$$

Continuamos este processo até J^n.

Voltamos para a nossa função $f(J = \alpha_0 J^0 + \alpha_1 J^1 + \ldots + \alpha_n J^n)$. A diagonal principal da matriz $\sum \alpha_i J^i$ é dada por

$$\alpha_0 + \alpha_1\lambda + \alpha_2\lambda^2 + \alpha_3\lambda^n + \ldots = f(\lambda) = \frac{f(\lambda)}{0!}.$$

Na primeira sobrediagonal temos:

$$\alpha_0.0 + \alpha_1.1 + \alpha_2.2\lambda + \alpha_3.3\lambda^2 + \ldots = \alpha_1 + 2\lambda\alpha_2 + 3\lambda^2\alpha_3 + \ldots = f'(\lambda) = \frac{f'(\lambda)}{1!}.$$

Na segunda sobrediagonal temos:

$$\alpha_0.0 + \alpha_1.0 + \alpha_2.1 + \alpha_3.3\lambda + \ldots = \alpha_2 + 3\lambda\alpha_3 = \frac{f''(\lambda)}{2} = \frac{f''(\lambda)}{2!}.$$

Em geral, na m-ésima sobrediagonal $1 \le m \le n$ é dado por

$$\frac{1}{m!}\sum_{i=m}^{\infty} i(i-1)\cdots(i-(m-1))\alpha_i\lambda^{i-m} = \frac{1}{m!}f^{(m)}(\lambda).$$

Assim, chegaremos à seguinte matriz:

$$f(J) = \begin{pmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \dots & \frac{f^n(\lambda)}{n!} \\ 0 & f(\lambda) & \ddots & \vdots \\ \vdots & & \ddots & \frac{f'(\lambda)}{1!} \\ 0 & 0 & \dots & f(\lambda) \end{pmatrix}.$$

Bibliografia

[1] BOLDRINI, José Luiz, et alli. *Álgebra Linear*, São Paulo: Editora Harper & Row do Brasil, 1980.

[2] CALLIOLI, Carlos A, et alli. *Álgebra Linear e Aplicações*, São Paulo: Editora Atual, 1987.

[3] CHATELIN, Françoise. *Eigenvalues of Matrices*, Inglaterra: Editora Wiley, 1993.

[4] GANTMACHER, F. R. *The theory of matrices*, New York: Chelsea Publishing Company, 1990.

[5] HOFFMAN, Kenneth & KUNZE Ray. *Álgebra Linear*, São Paulo: Editora Globo, 1970.

[6] SCHNEIDER, Hans & BARKER, George Phillip. *Matrices and Linear Algebra*, New York: Dover Publications, 1989.

[7] STRANG, Gilbert. *Linear Algebra and its applications*, Estados Unidos: Harcourt Brace & Company International, 1988.